Comparison of *Drosophila melanogaster* embryo and adult proteome by SWATH-MS reveals differential regulation of protein synthesis, degradation machinery and metabolism modules

Bertrand Fabre[1,2,4,5,†,*], Dagmara Korona[3,4,†], Jonathan G. Lees[6], Ikrame Lazar[5], Ido Livneh[5], Manon Brunet[1,2,4], Christine A. Orengo[6], Steven Russell[3,4], Kathryn S. Lilley[1,2,4,*]

1. Cambridge Centre for Proteomics, Department of Biochemistry, University of Cambridge, Cambridge, U.K

2. Department of Biochemistry, University of Cambridge, University of Cambridge, Cambridge, U.K

3. Department of Genetics, University of Cambridge, University of Cambridge, Cambridge, U.K

4. Cambridge Systems Biology Centre, University of Cambridge, Cambridge, U.K

5. Technion Integrated Cancer Center (TICC), The Rappaport Faculty of Medicine and Research Institute, Haifa, Israel

6. Institute of Structural and Molecular Biology, University College London, London, United Kingdom

† These authors contributed equally to this work.

* Correspondence:

Pr. Kathryn S. Lilley, Ph.D., Cambridge Centre for Proteomics, Cambridge Systems Biology Centre, Department of Biochemistry, University of Cambridge, Cambridge CB2 1QR, U.K. Telephone: 01223 760255; e-mail: k.s.lilley@bioc.cam.ac.uk

Dr. Bertrand Fabre, Ph.D., Technion Integrated Cancer Center (TICC), The Rappaport Faculty of Medicine and Research Institute, Haifa, Israel, e-mail: bertrand.fabre@cantab.net

ABSTRACT

An important area of modern biology consists in understanding the relationship between genotype and phenotype. However, to understand this relationship it is essential to investigate one of the principal links between them: the proteome. With the development of recent mass-spectrometry approaches it is now possible to quantify entire proteomes and thus relate them to different phenotypes. Here we present a comparison of the proteome of two extreme developmental states in the well-established model organism *Drosophila melanogaster*: adult and embryo. Protein modules such as ribosome, proteasome, tricarboxylic acid cycle, glycolysis or oxidative phosphorylation were found differentially expressed between the two developmental stages. Analysis of post-translation modifications of the proteins identified in this study indicates that they generally follow the same trend as their corresponding protein. Comparison between changes in the proteome and the transcriptome highlighted patterns of post-transcriptional regulation for the subunits of protein complexes such as the ribosome and the proteasome, whereas protein from modules such as TCA cycle, glycolysis and oxidative phosphorylation seem to be co-regulated at the transcriptional level. Finally, the impact of the endosymbiont *Wolbachia pipientis* on the proteome of both developmental states was also investigated.

Keywords:

SWATH-MS; Proteomics; Drosophila melanogaster; Embryo; Adult fly; Post-transcriptional regulation

INTRODUCTION

Proteins are responsible for most of the enzymatic reactions and the structural organization of cells and organisms [1,2]. It is now clear that different physiological state or stresses can affect the expression of specific proteins [3–5]. Although it is well known that the genome is largely stable across the life of an organism, the deployment of the proteome remains elusive [6–8]. Although many studies have linked molecular markers with particular phenotypes, notably through the identification of quantitative trait locus (QTL) [6,9], few have investigated changes in the proteome between, for example, different developmental states [10–13] and the relation between phenotype and proteome is still not well-defined [6,8]. To monitor changes in the proteome between two very different developmental states of an organism, quantitative mass-spectrometry is a particularly suitable method [6]. Recent MS approaches based on Data Independent Acquisition (DIA) methods, such as SWATH-MS, combine deep proteome coverage with ions selectivity similar to targeted proteomics, resulting in precise quantification of a large number of proteins [6,14]. With its small size and short life cycle, features that allow large cohorts of individuals to be rapidly collected, *Drosophila melanogaster* is a model organism particularly well-suited to study development and aging [15]. After egg laying each individual goes through a complete phenotypic metamorphosis within a few days, going through embryogenesis, larval life, metamorphosis in the pupa and the emergence of an adult in less than 9 days [15]. However, little information is known about the evolution of the proteome during this time and more particularly how different it can be between the two extremes of *Drosophila melanogaster* development, the embryo and the adult fly.

In the present study, we used the model organism *Drosophila melanogaster* to compare the proteome of two extreme phenotypic states, the embryo and the adult fly. The comparison of these two developmental states is of particular interest because it able to highlight features of terminally differentiated (adult) or differentiating (embryo) states of an organism. In addition, due to the comprehensive work of the modENCODE initiative, large data sets are available for RNA levels across different *Drosophila melanogaster* developmental states [16]. In addition to development, we also wished to determine whether quantitative differences in the proteome elicited by infection could be detected and to this end examined flies with and without the endosymbiont bacteria *Wolbachia pipientis* which is maternally inherited and present in over 60% of Drosophila strains [17].

Using SWATH-MS, we observed large changes in the proteome between embryos and adult flies (51% of the proteins quantified were differentially expressed) but no significant modulation of protein expression due to the presence of *Wolbachia pipientis*. The proteins differentially expressed between embryos and adult flies belong to protein modules such as ribosome, proteasome, protein processing in the ER, spliceosome, tricarboxylic acid (TCA) cycle, glycolysis and oxidative phosphorylation. Interestingly, the peptides bearing post-translation modifications (PTMs) identified and quantified in our analysis seem to follow the same trend as the corresponding protein. Finally, changes in the proteome between embryos and adult flies were compared with modENCODE transcriptome data [16], identifying patterns of post-transcriptional regulation for subunits of protein complexes including the ribosome, the proteasome and the eIF3 complex. In contrast, protein from modules such as TCA cycle, glycolysis and oxidative phosphorylation seem to be co-regulated at the transcriptional level.

MATERIAL AND METHODS

Drosophila husbandry, collection and protein extraction

D. melanogaster flies from the sequenced iso-1 strain (Bloomington Stock Centre: y1; Gr22biso-1 Gr22diso-1 cn1 CG33964iso-1 bw1 sp1; LysCiso-1 MstProxiso-1 GstD5iso-1 Rh61) were kept in a 12-h light-dark cycle at 25 °C and 75% relative humidity on standard yeast-cornmeal media. Embryos were collected from yeasted grape-juice agar plates, dechorionated with 50% bleach, washed with water, frozen in liquid nitrogen and kept at -80 °C. Embryos and adult flies samples were lysed as described in [18].

Wolbachia infected and cured D. melanogaster strains were a kind gift from Prof. FM. Jiggins (Cambridge, UK). The Wolbachia infection status of fly lines was checked by PCR (Polymerase Chain Reaction) using the diagnostic primers wsp81F and wsp691R [19]. DNA from flies was extracted using microLYSIS- Plus (microzone) according to the manufacturer's instructions. For the Wolbachia-infected lines, the wsp and 16S rRNA genes were amplified by PCR and checked by Sanger sequencing: 16Swol-F: 5'-TTGTAGCCTGCTATGGTATAACT-3'; 16SWol-R: 5'-GAATAGGTATGATTTTCATGT-3' [20]; wsp81-F: 5'-TGGTCCAATAAGTGATGAAGAAAC-3'; wsp691-R: 5'-AAAAATTAAACGCTACTCCA-3'.

Five replicates of 50 embryos/adults were performed for each condition.

Sample preparation for mass-spectrometry analysis

In gel digestion was used as the sample preparation method as described in [21].

Generation of the spectral library

The SWATH assay library used in this study is described in [22]. Post-translational modifications acetyl (N-ter), acetyl (K), GlyGly (K), methyl (KR) and Dimethyl (KR) were added in the MaxQuant search as variable modifications to generate the PTMs spectral libraries.

SWATH-MS acquisition

SWATH-MS data were acquired on a Sciex Triple-TOF 6600 as described in [12].

SWATH-MS data analysis

Spectronaut 10 (Biognosys) was used to analyse the SWATH data. Default settings were used except that a Q-value of 10^{-5} (corresponding to a FDR of 0.001% at the peptide level) was applied. Proteins with at least two peptides were used for quantitative analysis. The Top3 peptides method [23,24] was used for protein intensity. Up to 6 fragments were selected per peptide. The results were imported into R and Microsoft Excel for further analysis. For the comparison between adult and embryo, all the embryos samples (infected or not with Wolbachia) were compared to all the adult samples (infected or not with Wolbachia). To account for the multiplicity of null hypotheses being tested, the p-values from a Welch's t-test were adjusted using the Benjamini-Hochberg procedure [25]. STRING v10.5 [26] and FlyMine [27] were used for GO and KEGG analysis and network generation. For the comparison between the proteome and the transcriptome, modEncode mRNA expression data [16] were downloaded and mRNA expression indexes were calculated for each gene as the averaged expression values of all the embryo stages (from 0 to 24 hours after laying) and the averaged expression values of all the adult flies samples (head, carcass, digestive system, testes, accessory glands, ovaries, for male and female 1, 4, and 20 days post-eclosion). To account for possible effects on the dynamic range of abundances proteins or transcripts more statistical analysis were performed. To determine if the correlations for different protein modules were

different between the transcript and protein level, random sampling experiments were performed with the sample size matching the number of genes quantified in each protein module. For every protein module, a skew-normal distribution was fitted to the Spearman's rank correlation coefficient distribution of 1,000 random samples and the p-value was estimated. This type of analysis provides the significance of the correlation for each protein module, at the mRNA or protein level, compared to a random sampling and accounting for differences in dynamic range of abundances.

Data availability

All the mass spectrometry data have been deposited with the MassIVE repository with the dataset identifier: MSV000082812. In addition, a table including the quantitative analysis is provided in the supplementary information (Supplementary Table).

Western-Blot analysis

The antibodies against acetylated lysine (9441) and Rps6 (54D2) were purchased from Cell signalling Technology. The antibodies against Rpn8 (sc-390705), Rpt1 (sc-98691) and Histone H3 (sc-8654) were purchased from Santa Cruz Biotechnology. The antibody against Blw (ATP5a) (ab14748) was purchased from Abcam.

RESULTS AND DISCUSSION

SWATH-MS workflow to study the proteome of embryo and adult *Drosophila melanogaster*

In order to investigate changes of the proteome between embryo and adult fly, a SWATH-MS workflow was used (Figure 1A). SWATH-MS is a reliable MS method with performance at least

comparable with Data Dependent Acquisition (DDA) in terms of proteome coverage and precision [12,22,28]. Flies and embryos, infected with *Wolbachia pipientis* or cured (as confirmed by PCR, Figure S1), were lysed, the proteins were digested in gel with trypsin and the peptides were analysed on a TripleTOF-6600 (Sciex) in data independent acquisition mode (Figure 1A). The data were analysed with Spectronaut 10 (Biognosys) using a spectral library produced from embryo and adult fly samples [22] (Figure 1A). Using this workflow 1,533 proteins were quantified with at least 2 peptides in all the samples. Although the number of proteins quantified seems low compared to studies in other organisms [29,30], it is similar to a recent article where SWATH-MS was used to monitor the proteome of *Drosophila melanogaster* [31]. The number of proteins quantified is in part linked to the depth of coverage of the spectral library used to analyse the data. Indeed, many SWATH-MS data with good proteome coverage were performed in organisms where a large spectral library is available [32]. The median coefficient of variation (CV) observed for proteins between four injection replicates was 3% (Figure S2A) with very good linearity between the replicates (Figure S2B and C). The median CVs between biological replicates was higher, around 13% for both embryos and adult flies (Figure 1B and Figure S2A), which is consistent with other studies [12,33]. Interestingly, the variability across samples was higher (median CV of 45%) than the technical and biological variability (Figure S2A). When comparing the samples pairwise, a clear difference could be observed between the embryo and adult samples (Figure 1C and Figure S3). Indeed, Pearson correlation coefficient between adult or embryo samples were more than 0.9 whereas they were at best 0.68 between embryo and adult samples (Figure 1C and Figure S3).

Effect of *Wolbachia pipientis* on the proteome of embryo and adult *Drosophila melanogaster*

High Pearson correlation coefficients were observed between samples with or without *Wolbachia pipientis* (Figure 1C, Figure S3 and Figure S4A and B). Hierarchical cluster analysis showed that

samples with or without *Wolbachia pipientis* could not be differentiated with either embryos or adult flies (Figure 1D). When investigating the effect of *Wolbachia pipientis* on the proteomes of adult flies or embryos, no significant (FDR = 5%) changes was observed (Figure S4C and Figure S5). These results indicate that *Wolbachia pipientis* does not seem to profoundly affect the proteome of *Drosophila melanogaster* at the whole organism level for the stages examined. However, it is important to note that our study covers the most abundant proteins from the whole organism and it is possible that the analysis of low abundance proteins and/or tissue specific samples might identify differences in protein expression level as demonstrated in another study [34].

Comparison of *Drosophila melanogaster* embryo and adult proteomes

Although no significant changes were identified due to the presence of Wolbachia, the proteomes of the two different developmental states were dramatically different, with approximately 50% of the quantified proteins found to be differentially expressed between stages (Figure 2A). 407 proteins were significantly enriched in adults (log_2 fold change > 1 at a FDR of 1%) and 371 were significantly enriched in embryos (log_2 fold change < -1 at a FDR of 1%) (Figure 2A). To support our SWATH-MS data processing, an analysis was also performed using Skyline [35], and very similar results were obtained (Figure S6).

Different expression levels of protein synthesis and degradation machineries and metabolic pathways between embryo and adult flies

In order to identify differentially regulated pathways, a KEGG pathway analysis was performed. Protein complexes including the ribosome, proteasome and spliceosome were up regulated in the embryo compared to adult flies (Figure 2B and Figure S7A). Proteins involved in RNA transport, aminoacyl-tRNA biosynthesis and protein processing in endoplasmic reticulum were also more

abundant in the embryo (Figure 2B and Figure S7A). On the other hand, proteins related to metabolic pathways such as oxidative phosphorylation, TCA cycle and glycolysis were more abundant in adult flies (Figure 2C and Figure S7B).

Supporting the above observations, when plotting comparative abundances for the two developmental stages, protein components of the ribosome, eIF3 and proteasome complexes, all appeared to be more abundant in the embryo (Figure 3A-B-C). When plotting proteins involved in oxidative phosphorylation, TCA cycle and glycolysis, the same trend was observed in adults, with few exceptions, indicating that almost all the proteins of these pathways were more abundant in the adults (Figure 3D-E-F and Figure S8). Western-blot validations were performed for some of these proteins (Figure 3G) and are in agreement with the SWATH-MS data. Interestingly, histone expression was mostly unchanged between embryo and adult fly (Figure 3G and Figure S9).

Overall, these data suggest that protein synthesis and degradation machineries are more highly expressed in embryos whereas metabolic pathways are more abundant in adult flies (Figure 2, 3 and Figure S8). Embryonic development, which lasts around 24 hours, is a process which requires dynamic changes in protein expression to progress towards the development of a crawling larva [10,12]. High protein synthesis and degradation are likely to be required to facilitate such a rapid modulation of the proteome. Interestingly, an increase in metabolic enzyme mRNA level was observed during Drosophila embryogenesis [36] and also in rat [37] and mouse [38], where expression was higher in adult compared to embryo. However, it is not clear why proteins involved in oxidative phosphorylation, TCA cycle and glycolysis are more abundant in adult flies compared to embryos (Figure 3D-E-F and Figure S8), it is possible that high levels of such proteins might explain the higher resistance to anoxia shown by adult flies compared to less differentiated stages such as the larva [39].

Post-translational modifications follow the same trend as their corresponding proteins

To analyse post-translational modifications (PTMs), we generated a spectral library which includes acetylated, phosphorylated, methylated, dimethylated and ubiquitinated peptides (see Material and methods section for more details). 160 modified peptide precursors were quantified in all samples with only 2 bearing 2 different modifications (acetylation and phosphorylation) (Figure 4A and Supplementary Table). The most frequent PTM identified was acetylation (Figure 4A).

Surprisingly, when the modified peptide fold change was plotted against the corresponding protein fold change a clear correlation could be observed (r_s = 0.75, Figure 4B). This result indicates that, for the proteins and PTMs quantified in our study (which are the most abundant proteins from the embryo and adult fly), most PTMs follow the same expression trend as the other peptides (modified or not) of the corresponding protein. Thus, it seems unlikely that these proteins are differentially post-translationally modified between embryo and adult fly. This was most notably observed in the case of histones (Figure 4C-D), for example, the histone H1 N-terminus acetylated peptide ac-SDSAVATSASPVAAPPATVEK intensity strongly correlates with the intensity of the histone H1 protein (r_s = 0.91, Figure 4C). In addition, an acetylated lysine western blot showed three strong bands at molecular weights corresponding to histones and with similar intensities between embryo and adult samples (Figure 4D). When generating the PTMs spectral library, acetylated peptides could be identified for histone H1 and H3 (Figure S10A-B). It is well known that acetylation of histone is very important for chromatin remodelling and thus transcriptional activity, which is essential during embryogenesis as many transcriptional events happen during this process [40]. Our results (Figure 4C-D) suggest that there are no major differences in total histone acetylation between adult flies and embryos, suggesting that

modifications of histone acetylation probably happen on discrete sites on chromatin rather than global changes of histones acetylation profile during embryogenesis.

Interestingly, the peptide reporting K48 ubiquitin linkage (peptide with a remnant Gly-Gly linked to the lysine 48 of ubiquitin) could be identified and quantified in our data (Figure 4E, Figure S10C and Figure S11). This peptide was 2.4 times more abundant in the embryo compared to adult flies ($p = 7.4E-06$) (Figure 4E and Figure S11). This observation is particularly interesting since the K48 ubiquitin chain is considered to be one of the major signals for protein degradation by the 26S proteasome [41,42]. The high levels of K48 ubiquitin peptide detected in in embryo samples may be indicative of a high level of proteins tagged for proteasomal degradation. We suggest that embryos require higher protein degradation capacity as development unfolds as reflected by the higher expression of proteasome subunits compared to adults (Figure 2B and Figure 3A and G). This result is in agreement with the decrease in proteasome activity observed when human embryonic stem cells were differentiated into neurons [43].

Relationship between proteome and transcriptome

In order to explore the relationship between the transcriptome and the proteome, we used data from the modENCODE project [16]. When comparing protein and mRNA levels within both developmental stages, a positive correlation was observed (Figure S12A and B). The Spearman's rank correlation coefficient observed between protein and transcript level in adult flies ($r_s = 0.63$) was very similar to that describe in a previous study ($r_s = 0.66$) [44]. However, protein and mRNA expression levels were quite different between embryos and adult flies ($r_s = 0.59$ for both protein and mRNA level) (Figure 5A and B). The adult/embryo ratios were plotted for protein and mRNA and a positive correlation was observed (Figure 5C). The Spearman's rank correlation coefficient observed ($r_s = 0.52$) was similar to the one describe in a previous study investigating a Drosophila

brain tumor model (r_s = 0.61) [45]. When considering the genes quantified, 66% of the mRNAs and proteins showed consistent fold change directions with the remaining 34% showing fold changes in the opposite direction (Figure 5D). These observations suggest that for two thirds of the genes studied here when a mRNA is more abundant at one developmental stage it is also the case for the corresponding protein.

Protein and mRNA levels comparison highlights differential and co-regulation of protein modules

Next, the relation between protein expression and mRNA level was investigated for subunits of some protein complexes and protein modules (Figure 6). Interestingly, clear positive correlations were found between the proteins belonging to a common protein complex or protein module (Figure 6A, B, C, D and Figure S13A, B, C, D). However, while such correlations were observed for mRNA of protein modules linked with metabolism, such as TCA cycle (Figure 6H), glycolysis (Figure 6G) and oxidative phosphorylation (Figure S13H), there was far less correlation for mRNAs encoding subunits of protein complexes such as ribosome (Figure 6E), proteasome (Figure 6F), TCP-1 Ring Complex (TRiC) chaperonin (Figure S13E), ATP synthase (Figure S13F) and eIF3 (Figure S13G). To account for possible effects on the dynamic range of abundances of proteins or transcripts, random sampling experiments were performed with the sample size matching the number of genes quantified in each protein module. This analysis strengthens the significance of the measured correlation for each protein module, at the mRNA or protein level (Figure S14A-B). A Fisher r-to-z transformation was performed to assess the difference between the correlation observed at the protein and mRNA level for these protein modules. For most of the protein complexes studied here, strong z differences were observed between correlations measured for mRNA and proteins (Figure S14). Indeed, for ribosome (z_{diff} = 9.6, $p < 2.2E-16$), proteasome (z_{diff} =

6.54, p < 2.2E-16), eIF3 (z_{diff} = 3.38, p = 0.0007) and ATP synthase (z_{diff} = 3.38, p = 0.002) significant differences could be observed between mRNA and protein correlations. This was not the case for the TRiC chaperonin complex though (z_{diff} = 1.34, p = 0.1802). Concerning metabolism related protein modules, significant differences between mRNA and protein correlations could be observed only for oxidative phosphorylation (z_{diff} = 2.21, p = 0.0271). For TCA cycle and glycolysis, no significant differences could be observed (z_{diff} = -0.56, p = 0.5755 and z_{diff} = 1.65, p = 0.0989, respectively).

Taken together, these observations suggest patterns of post-transcriptional regulation for some protein complexes, most notably with the ribosome and proteasome (Figure 6A, B, E, F and Figure S14). Such patterns for protein complexes might be due to translation co-regulation of the subunits or complex stabilization when these macromolecular structures are assembled. These protein complexes are known to be particularly stable in eukaryotic cells (median half-life of 119, 130, 168, 42 and 72 hours for proteasome, ribosome, ATP synthase, eIF3 and TRiC, respectively in non-synchronized NIH3T3 mouse fibroblasts) [46]. On the other hand, metabolism related modules such as TCA cycle, glycolysis and oxidative phosphorylation seem to be co-regulated at the transcriptional level (Figure 6C, D, G, H, Figure S13D, H and Figure S14). Although, a significant difference could be observed between mRNA and protein correlations for oxidative phosphorylation, clear correlations were measured for both mRNA (r_s = 0.77) and protein (r_s = 0.89) between embryo and adult fly (Figure S13D, H and Figure S14), still suggesting a co-regulation of transcription of the corresponding genes.

CONCLUSIONS

In this study we compared the proteome of *Drosophila melanogaster* in two extreme developmental states, the embryo and the adult fly. Many differences in protein expression could be observed between the two different states, suggesting very different proteomes (Figure 2, Figure S7 and Figure S8). On the contrary, we did not detect any significant changes in protein expression due to the presence of *Wolbachia pipientis* in embryos or adult flies (Figure S5). Deeper or organ focused studies may bring to light effects of this endosymbiont on the *Drosophila* proteome, but it does not seem to affect expression of the most abundant proteins at the level of the whole organism.

The comparison between the embryo and adult fly proteomes revealed changes in expression of protein modules such as protein synthesis (ribosome and eIF3), protein degradation (proteasome) and metabolic pathways (TCA cycle, glycolysis and oxidative phosphorylation) (Figure 2 and 3). Interestingly, protein synthesis and degradation modules showed patterns of post-transcriptional regulation whereas metabolic related modules seem to be co-regulated through transcription (Figure 6, Figure S13 and Figure S14). This study highlights the importance of these protein modules across the life-cycle and suggests both transcriptional and post-transcriptional co-regulation mechanisms are used (Figure 6, Figure S13 and Figure S14). Further study will be necessary to understand how these protein modules impact the developmental progression of the organism.

Finally, we investigated how PTMs were distributed between embryo and adult fly (Figure 4) and were surprised to find that modified peptides generally followed the same trend as their corresponding proteins (Figure 4B, C and D). This suggests that PTMs do not appear to vary much between the two developmental stages, although we realise this conclusion is only applicable in the cases of the most abundant proteins and at the level of the whole animal. Regulation of PTMs

proteins in particular organs, at specific stages of development and for low abundant proteins, such as transcription factors, is surely of crucial importance for the development of an organism [47,48].

SUPPORTING INFORMATION:

The following supporting information is available free of charge at ACS website http://pubs.acs.org.

Figure S1. Verification of Wolbachia infection status of fly lines by PCR

Figure S2. Reproducibility of the SWATH-MS workflow

Figure S3. Pearson coefficient of correlation calculated between all the samples from pairwise comparison.

Figure S4. Comparison of Wolbachia infected and cured samples

Figure S5. Effect of Wolbachia pipientis on the embryo and adult fly proteomes

Figure S6. Skyline validation of the SWATH-MS data

Figure S7. Network representation of differentially regulated proteins between adult fly and embryo

Figure S8. Proteins from glycolysis and TCA cycle are more abundant in adult fly compared to embryo

Figure S9. Histones remain constant between embryo and adult fly

Figure S10. MSMS spectra for acetylated histones and K48 ubiquitin linkage

Figure S11. Extracted ion chromatograms of the peptide reporting K48 ubiquitin linkage LIFAGK[GG]QLEDGR

Figure S12. Protein/mRNA correlation in adult fly and embryo

Figure S13. Correlation of Protein and mRNA components of modules

Figure S14. Random sampling experiments with the sample size matching the number of genes quantified in each protein module

Figure S15. Differences between Protein and mRNA components of modules correlation

Table S1. Table containing the SWATH-MS data

Table S2. Table containing the protein and mRNA data

Table S3. Table containing the post-translational modifications data

ACKNOWLEDGEMENTS

B.F. is funded by Biotechnology and Biological Science Research Council (Ref: BB/L002817/1) and a long term EMBO fellow (ALTF 1204-2015) cofounded by Marie Curie Actions (LTFCOFUND2013, GA-2013-609409). D.K. is funded by Biotechnology and Biological Science Research Council (Ref: BB/L002817/1). I.L. is supported by the Foulkes Foundation Fellowship. The authors would like to thank Pr. Aaron Ciechanover and Pr. Ze'ev A. Ronai for providing reagents.

The authors have declared no conflict of interest.

References

(1) Ghosal, D.; Löwe, J. Collaborative Protein Filaments. *EMBO J.* **2015**, *34* (18), 2312–2320. https://doi.org/10.15252/embj.201591756.

(2) De la Fuente, I. M. Elements of the Cellular Metabolic Structure. *Front. Mol. Biosci.* **2015**, *2* (April). https://doi.org/10.3389/fmolb.2015.00016.

(3) Ruzycki, P. A.; Tran, N. M.; Kefalov, V. J.; Kolesnikov, A. V.; Chen, S. Graded Gene Expression Changes Determine Phenotype Severity in Mouse Models of CRX-Associated

Retinopathies. *Genome Biol.* **2015**, *16* (1), 1–22. https://doi.org/10.1186/s13059-015-0732-z.

(4) Chen, E. H.; Hou, Q. L.; Wei, D. D.; Jiang, H. B.; Wang, J. J. Phenotypes, Antioxidant Responses, and Gene Expression Changes Accompanying a Sugar-Only Diet in Bactrocera Dorsalis (Hendel) (Diptera: Tephritidae). *BMC Evol. Biol.* **2017**, *17* (1), 1–14. https://doi.org/10.1186/s12862-017-1045-5.

(5) Vu, V.; Verster, A. J.; Schertzberg, M.; Chuluunbaatar, T.; Spensley, M.; Pajkic, D.; Hart, G. T.; Moffat, J.; Fraser, A. G. Natural Variation in Gene Expression Modulates the Severity of Mutant Phenotypes. *Cell* **2015**, *162* (2), 391–402. https://doi.org/10.1016/j.cell.2015.06.037.

(6) Aebersold, R.; Mann, M. Mass-Spectrometric Exploration of Proteome Structure and Function. *Nature* **2016**, *537* (7620), 347–355. https://doi.org/10.1038/nature19949.

(7) Muers, M. Gene Expression: Transcriptome to Proteome and Back to Genome. *Nat. Rev. Genet.* **2011**, *12* (8), 518. https://doi.org/10.1038/nrg3037.

(8) Liu, Y.; Beyer, A.; Aebersold, R. On the Dependency of Cellular Protein Levels on MRNA Abundance. *Cell* **2016**, *165* (3), 535–550. https://doi.org/10.1016/j.cell.2016.03.014.

(9) Williams, E. G.; Auwerx, J. The Convergence of Systems and Reductionist Approaches in Complex Trait Analysis. *Cell* **2015**, *162* (1), 23–32. https://doi.org/10.1016/j.cell.2015.06.024.

(10) Casas-Vila, N.; Bluhm, A.; Sayols, S.; Dinges, N.; Dejung, M.; Altenhein, T.; Kappei, D.; Altenhein, B.; Roignant, J.-Y.; Butter, F. The Developmental Proteome OfDrosophila Melanogaster. *Genome Res.* **2017**, *27* (7), 1273–1285. https://doi.org/10.1101/gr.213694.116.

(11) Jin, S.; Sun, D.; Song, D.; Wang, N.; Fu, H.; Ji, F.; Zhang, Y. ITRAQ-Based Quantitative

Proteomic Analysis of Embryonic Developmental Stages in Amur Sturgeon , Acipenser Schrenckii. *Sci. Rep.* **2018**, No. March 2017, 1–12. https://doi.org/10.1038/s41598-018-24562-1.

(12) Fabre, B.; Korona, D.; Groen, A.; Vowinckel, J.; Gatto, L.; Deery, M. J.; Ralser, M.; Russell, S.; Lilley, K. S. Analysis of Drosophila Melanogaster Proteome Dynamics during Embryonic Development by a Combination of Label-Free Proteomics Approaches. *Proteomics* **2016**, *16* (15–16), 2068–2080. https://doi.org/10.1002/pmic.201500482.

(13) Bamberger, C.; Martínez-Bartolomé, S.; Montgomery, M.; Lavallée-Adam, M.; Yates, J. R. Increased Proteomic Complexity in Drosophila Hybrids during Development. *Sci. Adv.* **2018**, *4* (2). https://doi.org/10.1126/sciadv.aao3424.

(14) Vowinckel, J.; Zelezniak, A.; Bruderer, R.; Mülleder, M.; Reiter, L.; Ralser, M. Cost-Effective Generation of Precise Label-Free Quantitative Proteomes in High-Throughput by MicroLC and Data-Independent Acquisition. *Sci. Rep.* **2018**, *8* (1), 1–10. https://doi.org/10.1038/s41598-018-22610-4.

(15) Riedel, G. F. Drosophila Melanogaster as an Experimental Organism. *Science (80-.).* **1985**, *7* (6), 191–204.

(16) Graveley, B. R.; Brooks, A. N.; Carlson, J. W.; Duff, M. O.; Landolin, J. M.; Yang, L.; Artieri, C. G.; Van Baren, M. J.; Boley, N.; Booth, B. W.; et al. The Developmental Transcriptome of Drosophila Melanogaster. *Nature* **2011**, *471* (7339), 473–479. https://doi.org/10.1038/nature09715.

(17) Richardson, M. F.; Weinert, L. A.; Welch, J. J.; Linheiro, R. S.; Magwire, M. M.; Jiggins, F. M.; Bergman, C. M. Population Genomics of the Wolbachia Endosymbiont in Drosophila Melanogaster. *PLoS Genet.* **2012**, *8* (12). https://doi.org/10.1371/journal.pgen.1003129.

(18) Fabre, B.; Korona, D.; Nightingale, D. J. H.; Russell, S.; Lilley, K. S. SWATH-MS Dataset of

Heat-Shock Treated Drosophila Melanogaster Embryos. *Data Br.* **2016**, *9*. https://doi.org/10.1016/j.dib.2016.11.028.

(19) Zhou, W.; Rousset, F.; O'Neill, S. Phylogeny and PCR-Based Classification of Wolbachia Strains Using Wsp Gene Sequences. *Proc. R. Soc. B Biol. Sci.* **1998**, *265* (1395), 509–515. https://doi.org/10.1098/rspb.1998.0324.

(20) O'Neill, S. L.; Giordano, R.; Colbert, A. M.; Karr, T. L.; Robertson, H. M. 16S RRNA Phylogenetic Analysis of the Bacterial Endosymbionts Associated with Cytoplasmic Incompatibility in Insects. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (7), 2699–2702. https://doi.org/10.1073/pnas.89.7.2699.

(21) Fabre, B.; Korona, D.; Nightingale, D. J. H.; Russell, S.; Lilley, K. S. SWATH-MS Data of Drosophila Melanogaster Proteome Dynamics during Embryogenesis. *Data Br.* **2016**, *9*. https://doi.org/10.1016/j.dib.2016.10.009.

(22) Fabre, B.; Korona, D.; Mata, C. I.; Parsons, H. T.; Deery, M. J.; Hertog, M. L. A. T. M.; Nicolaï, B. M.; Russell, S.; Lilley, K. S. Spectral Libraries for SWATH-MS Assays for Drosophila Melanogaster and Solanum Lycopersicum. *Proteomics* **2017**, *17* (21). https://doi.org/10.1002/pmic.201700216.

(23) Fabre, B.; Lambour, T.; Bouyssié, D.; Menneteau, T.; Monsarrat, B.; Burlet-Schiltz, O.; Bousquet-Dubouch, M. P. Comparison of Label-Free Quantification Methods for the Determination of Protein Complexes Subunits Stoichiometry. *EuPA Open Proteomics* **2014**, *4*, 82–86. https://doi.org/10.1016/j.euprot.2014.06.001.

(24) Schubert, O. T.; Ludwig, C.; Kogadeeva, M.; Zimmermann, M.; Rosenberger, G.; Gengenbacher, M.; Gillet, L. C.; Collins, B. C.; Rost, H. L.; Kaufmann, S. H. E.; et al. Absolute Proteome Composition and Dynamics during Dormancy and Resuscitation of Mycobacterium Tuberculosis. *Cell Host Microbe* **2015**, *18* (1), 96–108.

https://doi.org/10.1016/j.chom.2015.06.001.

(25) Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society*. 1995, pp 289–300. https://doi.org/10.2307/2346101.

(26) Szklarczyk, D.; Morris, J. H.; Cook, H.; Kuhn, M.; Wyder, S.; Simonovic, M.; Santos, A.; Doncheva, N. T.; Roth, A.; Bork, P.; et al. The STRING Database in 2017: Quality-Controlled Protein-Protein Association Networks, Made Broadly Accessible. *Nucleic Acids Res.* **2017**, *45* (D1), D362–D368. https://doi.org/10.1093/nar/gkw937.

(27) Lyne, R.; Smith, R.; Rutherford, K.; Wakeling, M.; Varley, A.; Guillier, F.; Janssens, H.; Ji, W.; Mclaren, P.; North, P.; et al. FlyMine: An Integrated Database for Drosophila and Anopheles Genomics. *Genome Biol.* **2007**, *8* (7). https://doi.org/10.1186/gb-2007-8-7-r129.

(28) Vowinckel, J.; Capuano, F.; Campbell, K.; Deery, M. J.; Lilley, K. S.; Ralser, M. The Beauty of Being (Label) -Free : Sample Preparation Methods for SWATH-MS and next-Generation Targeted Proteomics [v1 ; Ref Status : Awaiting Peer Review , Http://F1000r.Es/2ed]. *F1000Research* **2013**, *272* (0), 1–23. https://doi.org/10.12688/f1000research.2-272.v1.

(29) Liu, Y.; Borel, C.; Li, L.; Müller, T.; Williams, E. G.; Germain, P.-L.; Buljan, M.; Sajic, T.; Boersema, P. J.; Shao, W.; et al. Systematic Proteome and Proteostasis Profiling in Human Trisomy 21 Fibroblast Cells. *Nat. Commun.* **2017**, *8* (1), 1212. https://doi.org/10.1038/s41467-017-01422-6.

(30) Guo, T.; Li, L.; Zhong, Q.; Rupp, N. J.; Charmpi, K.; Wong, C. E.; Wagner, U.; Rueschoff, J. H.; Jochum, W.; Fankhauser, C. D.; et al. Multi-Region Proteome Analysis Quantifies Spatial Heterogeneity of Prostate Tissue Biomarkers. *Life Sci. Alliance* **2018**, *1* (2), e201800042. https://doi.org/10.26508/lsa.201800042.

(31) Okada, H.; Ebhardt, H. A.; Vonesch, S. C.; Aebersold, R.; Hafen, E. Proteome-Wide Association Studies Identify Biochemical Modules Associated with a Wing-Size Phenotype in Drosophila Melanogaster. *Nat. Commun.* **2016**, *7*, 12649.

(32) Rosenberger, G.; Koh, C. C.; Guo, T.; Röst, H. L.; Kouvonen, P.; Collins, B. C.; Heusel, M.; Liu, Y.; Caron, E.; Vichalkovski, A.; et al. A Repository of Assays to Quantify 10,000 Human Proteins by SWATH-MS. *Sci. Data* **2014**, *1*, 140031.

(33) Selevsek, N.; Chang, C.-Y.; Gillet, L. C.; Navarro, P.; Bernhardt, O. M.; Reiter, L.; Cheng, L.-Y.; Vitek, O.; Aebersold, R. Reproducible and Consistent Quantification of the *Saccharomyces Cerevisiae* Proteome by SWATH-Mass Spectrometry. *Mol. Cell. Proteomics* **2015**, *14* (3), 739–749. https://doi.org/10.1074/mcp.M113.035550.

(34) Christensen, S.; Pï¿½rez Dulzaides, R.; Hedrick, V. E.; Momtaz, A. J. M. Z.; Nakayasu, E. S.; Paul, L. N.; Serbus, L. R. Wolbachia Endosymbionts Modify Drosophila Ovary Protein Levels in a Context-Dependent Manner. *Appl. Environ. Microbiol.* **2016**, *82* (17), 5354–5363. https://doi.org/10.1128/AEM.01255-16.

(35) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. Skyline: An Open Source Document Editor for Creating and Analyzing Targeted Proteomics Experiments. *Bioinformatics* **2010**, *26* (7), 966–968. https://doi.org/10.1093/bioinformatics/btq054.

(36) Tennessen, J. M.; Baker, K. D.; Lam, G.; Evans, J.; Thummel, C. S. The Drosophila Estrogen-Related Receptor Directs a Metabolic Switch That Supports Developmental Growth. *Cell Metab.* **2011**, *13* (2), 139–148. https://doi.org/10.1016/j.cmet.2011.01.005.

(37) Nagao, M.; Parimoo, B.; Tanaka, K. Developmental, Nutritional, and Hormonal Regulation of Tissue-Specific Expression of the Genes Encoding Various Acyl-CoA Dehydrogenases and α- Subunit of Electron Transfer Flavoprotein in Rat. *J. Biol. Chem.* **1993**, *268* (32), 24114–

24124.

(38) Disch, D. L.; Rader, T. A.; Cresci, S.; Leone, T. C.; Barger, P. M.; Vega, R.; Wood, P. A.; Kelly, D. P. Transcriptional Control of a Nuclear Gene Encoding a Mitochondrial Fatty Acid Oxidation Enzyme in Transgenic Mice: Role for Nuclear Receptors in Cardiac and Brown Adipose Expression. *Mol. Cell. Biol.* **1996**, *16* (8), 4043–4051.

(39) Callier, V.; Hand, S. C.; Campbell, J. B.; Biddulph, T.; Harrison, J. F. Developmental Changes in Hypoxic Exposure and Responses to Anoxia in Drosophila Melanogaster. *J. Exp. Biol.* **2015**, *218* (18), 2927–2934. https://doi.org/10.1242/jeb.125849.

(40) Mannervik, M. Control of Drosophila Embryo Patterning by Transcriptional Co-Regulators. *Exp. Cell Res.* **2014**, *321* (1), 47–57. https://doi.org/10.1016/j.yexcr.2013.10.010.

(41) Grice, G. L.; Nathan, J. A. The Recognition of Ubiquitinated Proteins by the Proteasome. *Cell. Mol. Life Sci.* **2016**, *73* (18), 3497–3506. https://doi.org/10.1007/s00018-016-2255-5.

(42) Bousquet-Dubouch, M. P.; Fabre, B.; Monsarrat, B.; Burlet-Schiltz, O. Proteomics to Study the Diversity and Dynamics of Proteasome Complexes: From Fundamentals to the Clinic. *Expert Rev. Proteomics* **2011**, *8* (4), 459–481. https://doi.org/10.1586/epr.11.41.

(43) Vilchez, D.; Boyer, L.; Morantte, I.; Lutz, M.; Merkwirth, C.; Joyce, D.; Spencer, B.; Page, L.; Masliah, E.; Berggren, W. T.; et al. Increased Proteasome Activity in Human Embryonic Stem Cells Is Regulated by PSMD11. *Nature* **2012**, *489*, 304.

(44) Schrimpf, S. P.; Weiss, M.; Reiter, L.; Ahrens, C. H.; Jovanovic, M.; Malmström, J.; Brunner, E.; Mohanty, S.; Lercher, M. J.; Hunziker, P. E.; et al. Comparative Functional Analysis of the Caenorhabditis Elegans and Drosophila Melanogaster Proteomes. *PLoS Biol.* **2009**, *7* (3), 0616–0627. https://doi.org/10.1371/journal.pbio.1000048.

(45) Juschke, C.; Dohnal, I.; Pichler, P.; Harzer, H.; Swart, R.; Ammerer, G.; Mechtler, K.; Knoblich, J. A. Transcriptome and Proteome Quantification of a Tumor Model Provides

Novel Insights into Post-Transcriptional Gene Regulation. *Genome Biol.* **2013**, *14* (11), r133. https://doi.org/10.1186/gb-2013-14-11-r133.

(46) Schwanhüusser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M. Global Quantification of Mammalian Gene Expression Control. *Nature* **2011**, *473* (7347), 337–342. https://doi.org/10.1038/nature10098.

(47) Krause, H. M.; Gehring, W. J. Stage-Specific Phosphorylation of the Fushi Tarazu Protein during Drosophila Development. *EMBO J.* **1989**, *8* (4), 1197–1204. https://doi.org/10.1002/j.1460-2075.1989.tb03492.x.

(48) Akhmetova, K.; Balasov, M.; Svitin, A.; Chesnokova, E.; Renfrow, M.; Chesnokov, I. Phosphorylation of Pnut in the Early Stages of Drosophila Embryo Development Affects Association of the Septin Complex with the Membrane and Is Important for Viability. *G3:Genes|Genomes|Genetics; Genes|Genomes|Genetics* **2018**, *8* (1), 27–38. https://doi.org/10.1534/g3.117.300186.

Authors contributions:

BF, DK, SR and KSL designed experiments. BF, DK, IL, IL and MB performed experiments. BF and JGL analysed data. BF, CAO, SR and KSL supervised experiments. BF wrote the main manuscript. All authors reviewed and edited the manuscript.

Figure Legends:

Figure 1: SWATH-MS workflow used to study the proteome of Drosophila melanogaster embryo and adult fly, infected with Wolbachia pipientis or cured

A. Embryos and adult flies, infected with *Wolbachia pipientis* or cured (n = 5 for each condition), were lysed and proteins were extracted and digested in gel with trypsin. Peptides were analysed

on a Sciex Triple-TOF 6600 in SWATH acquisition mode. A spectral library produced in another study[22] was used to analysed the data with Spectronaut. B. Coefficients of variation were calculated for each protein quantified between all the biological replicates for each sample and for all the adult and embryo samples. C. Pearson coefficient of correlation was calculated pairwise between all the samples analysed. D. Hierarchical clustering of the samples analysed based on the measured proteins intensities.

Figure 2: Comparison of the embryo and adult fly proteome

A. Volcano plot representing the $\log_2$ ratio (adult/embryo) for each protein quantified and the corresponding Benjamini-Hochberg corrected p-value. The blue, red and grey dots represent the proteins more abundant in the embryo, more abundant in the adult fly and not differentially expressed between the two developmental stages, respectively. Dashed lines represent the 2 fold change for proteins up- and down-regulated and the 1% FDR threshold. B-C. Results from KEGG pathway analysis for the proteins more abundant in embryos (B) and more abundant in adult flies (C). The number of proteins identified in each pathway (protein count) and the p-value ($-\log_{10}$ transformed) are represented on the graphs.

Figure 3: Differential expression of protein modules between embryo and adult fly

A-F. Volcano plot representing the $\log_2$ ratio (adult/embryo) for each protein quantified and the corresponding Benjamini-Hochberg corrected p-value. The proteins belonging to proteasome (A), ribosome (B), eIF3 (C), glycolysis (D), oxidative phosphorylation (E) and TCA cycle (F) are represented of the graphs. G. Validation of the SWATH-MS by western-blot. The ratios measured by SWATH-MS are indicated on the right of the western-blot of the corresponding protein.

Histone H3 was used as a loading control as it was not found significantly differentially expressed in the SWATH-MS analysis.

Figure 4: Analysis of post-translational modifications in embryo and adult fly

A. Pie chart of the post-translational modifications (PTMs) (precursors) quantified in this study. B. Graph representing the log_2 fold change (adult/embryo) for the PTMs identified and their corresponding proteins. Each PTMs are labelled with a specific colour. The Spearman's rank correlation coefficient between all the quantified PTMs ratio and their corresponding protein is indicated on the graph. C. Graph representing the log_2 Intensity of the acetylated peptide ac-SDSAVATSASPVAAPPATVEK of the histone H1 plotted against the log_2 Intensity of the histone H1 protein. The resulting Spearman's rank correlation coefficient is indicated on the graph. D. Western blot analysis of adult fly and embryo samples using acetylated lysine and histone H3 antibodies. E. Abundance (Log_2 transformed peptide intensity) of the peptide reporting K48 ubiquitin linkage (LIFAGK[GG]QLEDGR) in embryo and adult fly.

Figure 5: Transcriptome and proteome relationship

A. Graph representing the embryo log_{10} protein intensities (calculated using the Top3 method) against the adult log_{10} protein intensities. The resulting Spearman's rank correlation coefficient is indicated on the graph. B. Graph representing the embryo log_{10} mRNA counts (calculated using the modENCODE data) against the adult log_{10} mRNA counts. The resulting Spearman's rank correlation coefficient is indicated on the graph. C. Graph representing the log_2 fold change (adult/embryo) for the proteins identified and their corresponding mRNA. The resulting Spearman's rank correlation coefficient is indicated on the graph. D. Graph representing number

of protein/mRNA counts based on the direction of the fold change (adult/embryo) for the proteins identified and their corresponding mRNA.

Figure 6: Correlation of Protein and mRNA components of modules

A-D. Graph representing the embryo $\log_{10}$ protein intensities (calculated using the Top3 method) against the adult $\log_{10}$ protein intensities. Ribosome (A), proteasome (B), glycolysis (C) and TCA cycle (D) components proteins intensities are represented on the graphs. E-H. Graph representing the embryo $\log_{10}$ mRNA counts (calculated using the modENCODE data) against the adult $\log_{10}$ mRNA counts. Ribosome (E), proteasome (F), glycolysis (G) and TCA cycle (H) components mRNA counts are represented on the graphs. The resulting Spearman's rank correlation coefficient is indicated on the graphs.

Figure 1

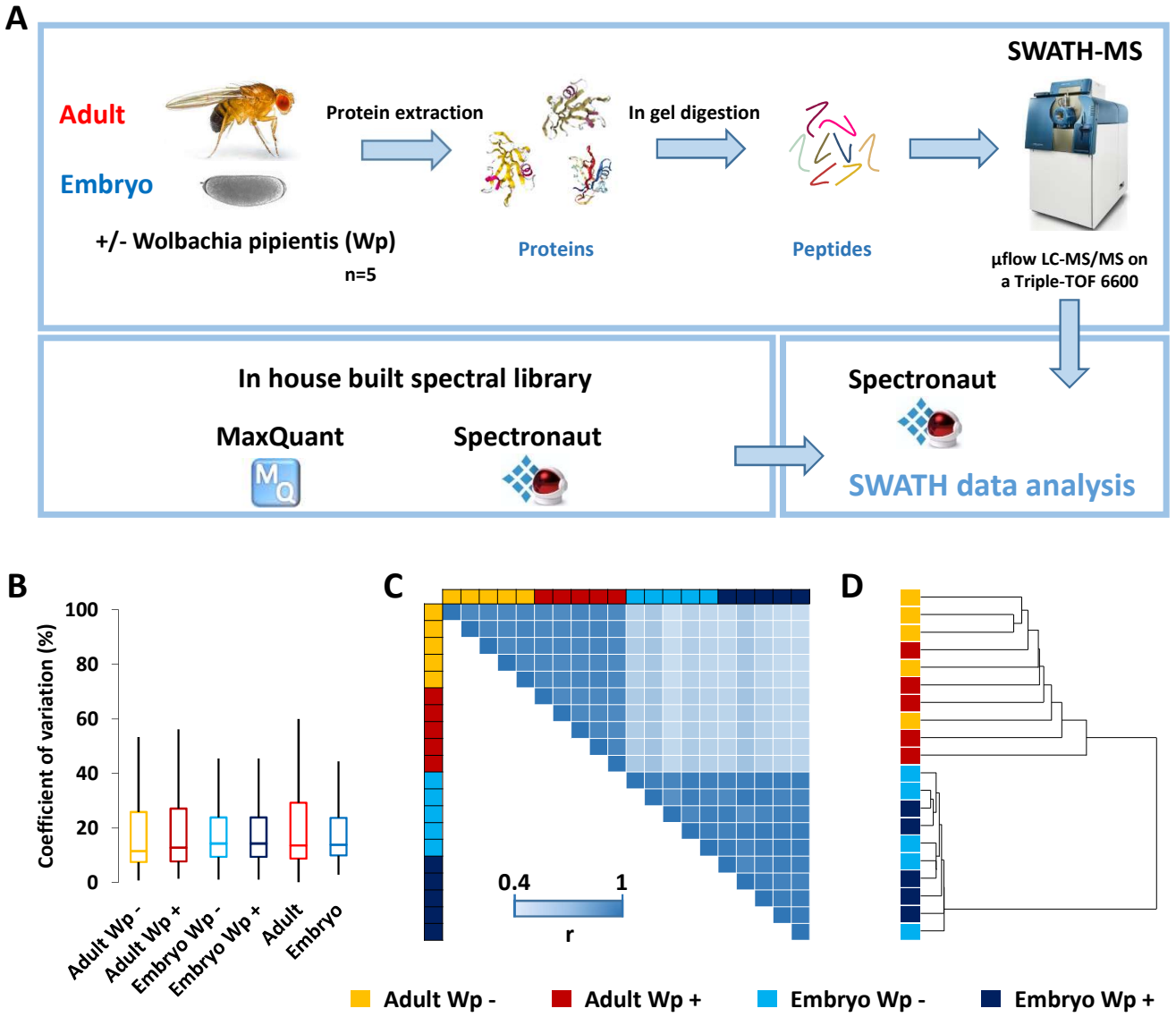

A
SWATH-MS
Adult
Embryo
Protein extraction
In gel digestion
+/- Wolbachia pipientis (Wp)
n=5
Proteins
Peptides
µflow LC-MS/MS on a Triple-TOF 6600
In house built spectral library
MaxQuant
Spectronaut
Spectronaut
SWATH data analysis
B
Coefficient of variation (%)
100
80
60
40
20
0
Adult Wp -
Adult Wp +
Embryo Wp -
Embryo Wp +
Adult
Embryo
C
0.4
1
r
D
Adult Wp -
Adult Wp +
Embryo Wp -
Embryo Wp +

Figure 2

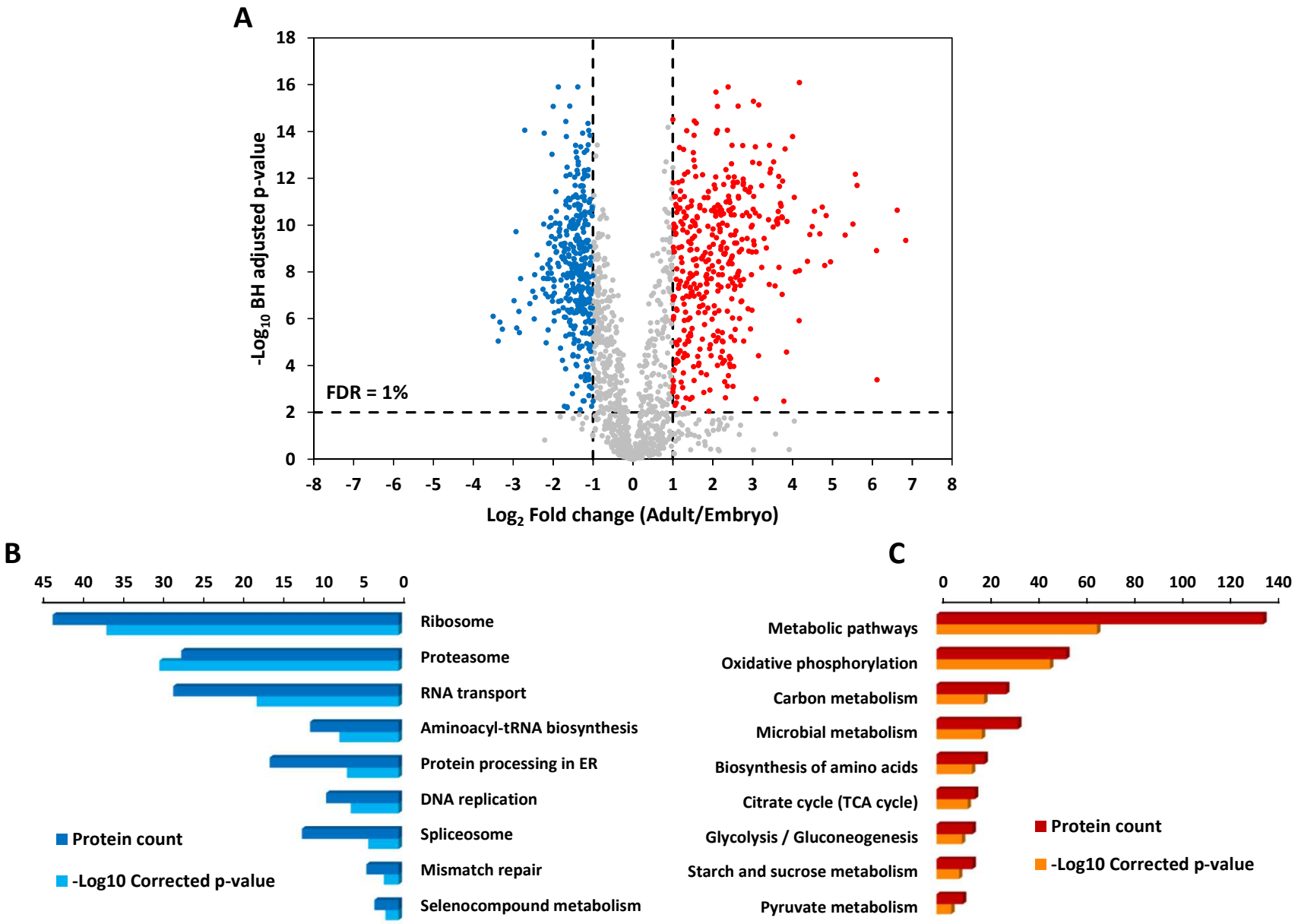

A
-Log10 BH adjusted p-value
Log2 Fold change (Adult/Embryo)
FDR = 1%
B
Ribosome
Proteasome
RNA transport
Aminoacyl-tRNA biosynthesis
Protein processing in ER
DNA replication
Spliceosome
Mismatch repair
Selenocompound metabolism
Protein count
-Log10 Corrected p-value
C
Metabolic pathways
Oxidative phosphorylation
Carbon metabolism
Microbial metabolism
Biosynthesis of amino acids
Citrate cycle (TCA cycle)
Glycolysis / Gluconeogenesis
Starch and sucrose metabolism
Pyruvate metabolism
Protein count
-Log10 Corrected p-value

Figure 3

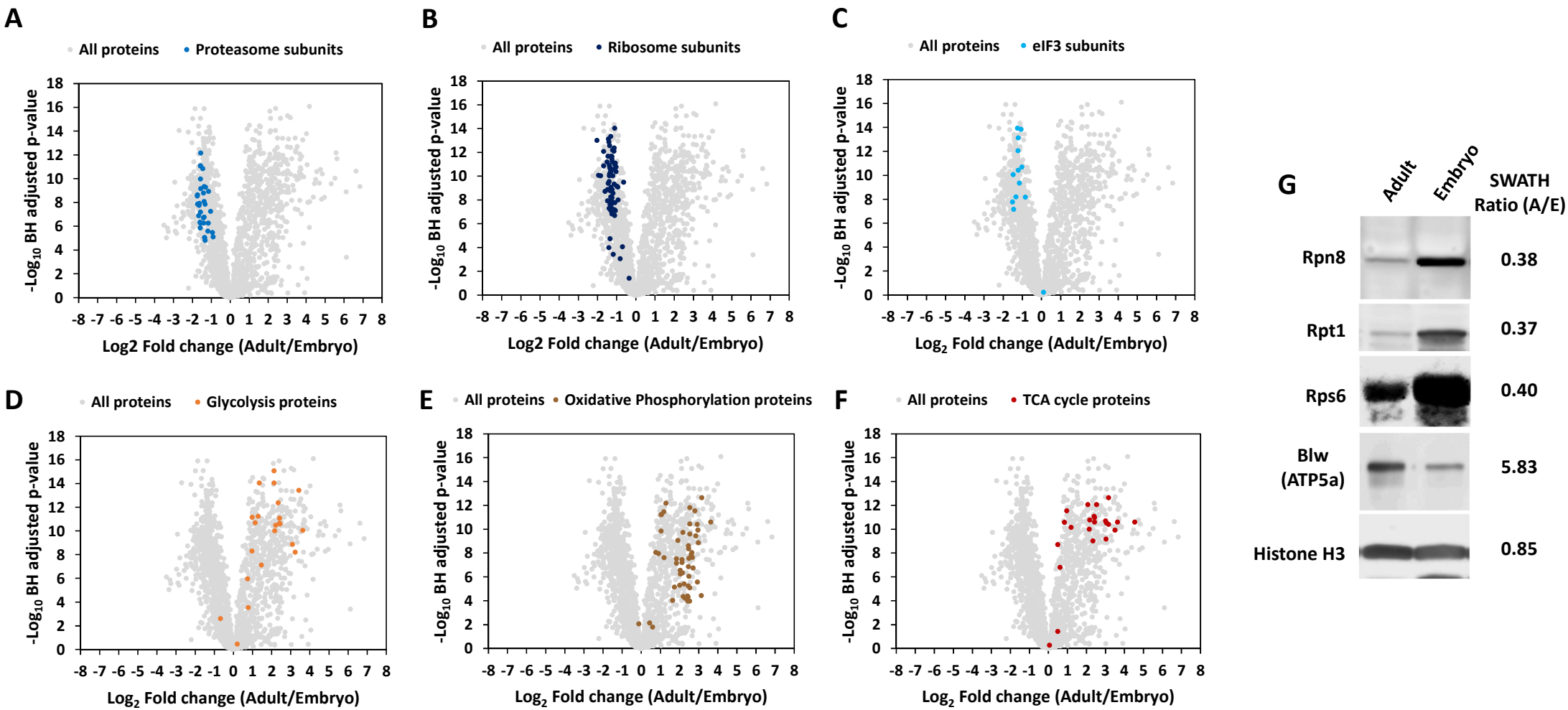

A
All proteins
Proteasome subunits
-Log10 BH adjusted p-value
Log2 Fold change (Adult/Embryo)
B
All proteins
Ribosome subunits
-Log10 BH adjusted p-value
Log2 Fold change (Adult/Embryo)
C
All proteins
eIF3 subunits
-Log10 BH adjusted p-value
Log2 Fold change (Adult/Embryo)
D
All proteins
Glycolysis proteins
-Log10 BH adjusted p-value
Log2 Fold change (Adult/Embryo)
E
All proteins
Oxidative Phosphorylation proteins
-Log10 BH adjusted p-value
Log2 Fold change (Adult/Embryo)
F
All proteins
TCA cycle proteins
-Log10 BH adjusted p-value
Log2 Fold change (Adult/Embryo)
G
Adult
Embryo
SWATH Ratio (A/E)
Rpn8
0.38
Rpt1
0.37
Rps6
0.40
Blw (ATP5a)
5.83
Histone H3
0.85

Figure 4

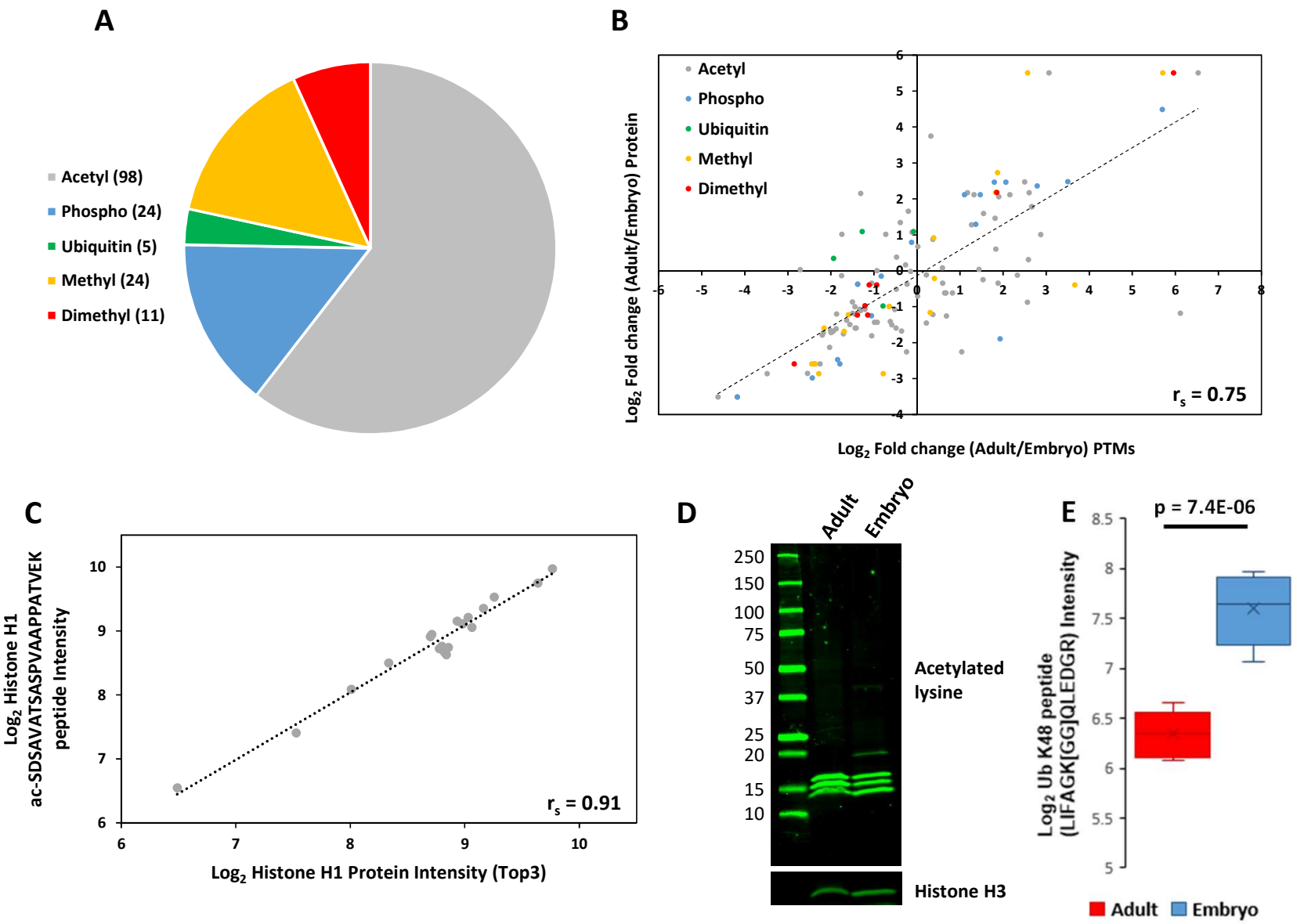

A
Acetyl (98)
Phospho (24)
Ubiquitin (5)
Methyl (24)
Dimethyl (11)
B
Acetyl
Phospho
Ubiquitin
Methyl
Dimethyl
Log2 Fold change (Adult/Embryo) Protein
Log2 Fold change (Adult/Embryo) PTMs
rs = 0.75
C
Log2 Histone H1 ac-SDSAVATSASPVAAPPATVEK peptide Intensity
Log2 Histone H1 Protein Intensity (Top3)
rs = 0.91
D
Adult
Embryo
250
150
100
75
50
37
25
20
15
10
Acetylated lysine
Histone H3
E
p = 7.4E-06
Log2 Ub K48 peptide (LIFAGK[GG]QLEDGR) Intensity
Adult
Embryo

Figure 5

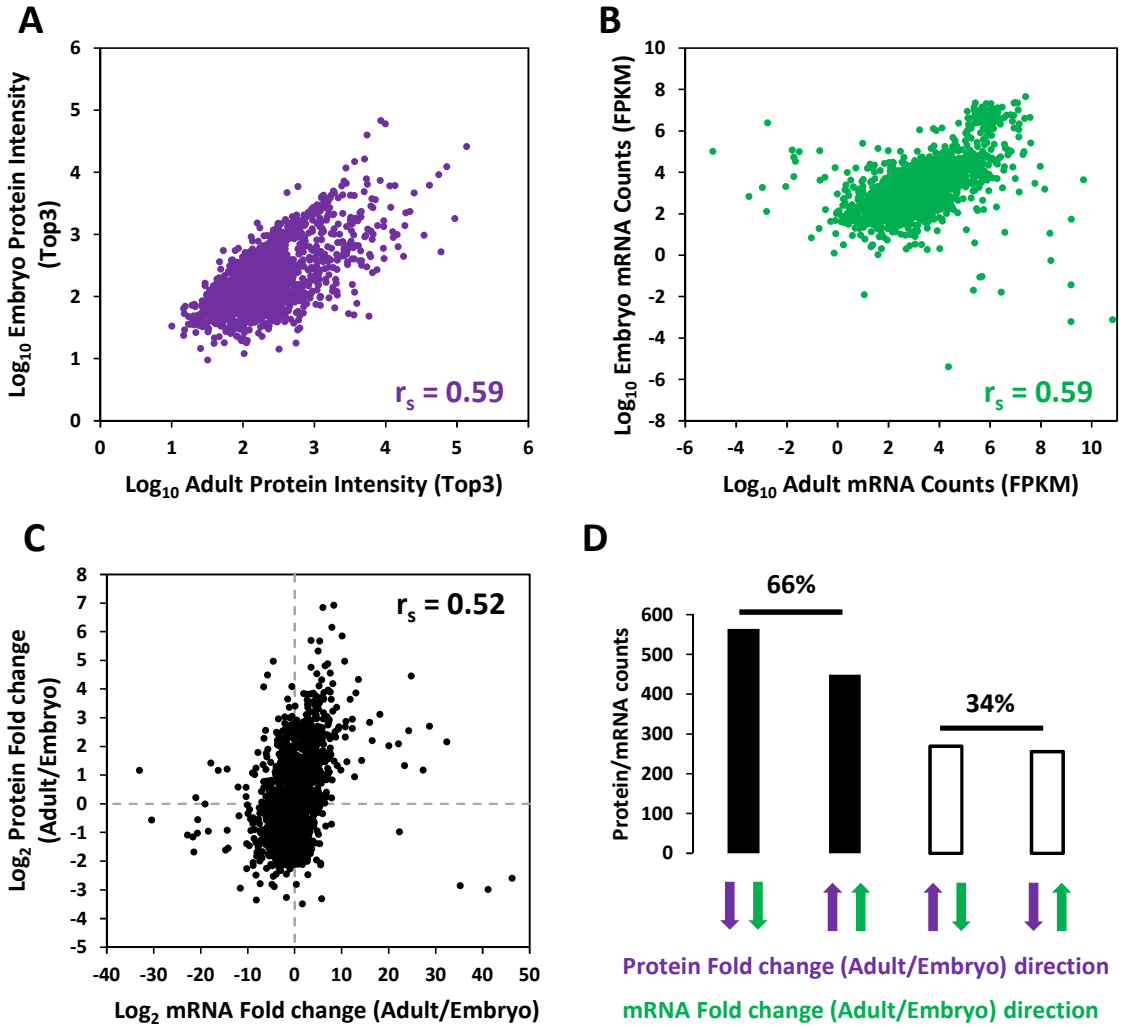
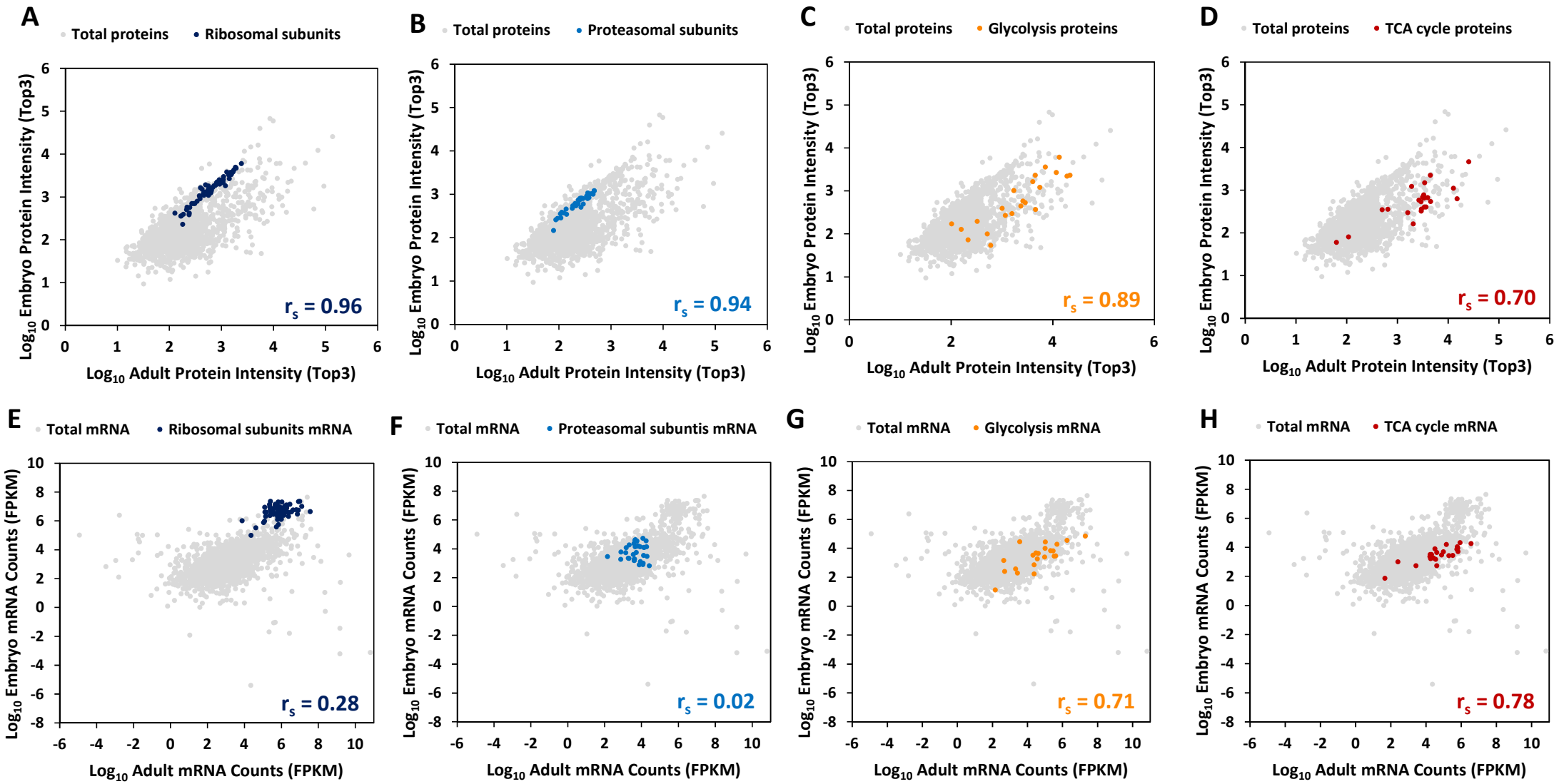

A
Log10 Embryo Protein Intensity (Top3)
Log10 Adult Protein Intensity (Top3)
rs = 0.59
B
Log10 Embryo mRNA Counts (FPKM)
Log10 Adult mRNA Counts (FPKM)
rs = 0.59
C
Log2 Protein Fold change (Adult/Embryo)
Log2 mRNA Fold change (Adult/Embryo)
rs = 0.52
D
Protein/mRNA counts
66%
34%
Protein Fold change (Adult/Embryo) direction
mRNA Fold change (Adult/Embryo) direction

Figure 6


A
Total proteins
Ribosomal subunits
Log10 Embryo Protein Intensity (Top3)
Log10 Adult Protein Intensity (Top3)
rs = 0.96
B
Total proteins
Proteasomal subunits
Log10 Embryo Protein Intensity (Top3)
Log10 Adult Protein Intensity (Top3)
rs = 0.94
C
Total proteins
Glycolysis proteins
Log10 Embryo Protein Intensity (Top3)
Log10 Adult Protein Intensity (Top3)
rs = 0.89
D
Total proteins
TCA cycle proteins
Log10 Embryo Protein Intensity (Top3)
Log10 Adult Protein Intensity (Top3)
rs = 0.70
E
Total mRNA
Ribosomal subunits mRNA
Log10 Embryo mRNA Counts (FPKM)
Log10 Adult mRNA Counts (FPKM)
rs = 0.28
F
Total mRNA
Proteasomal subuntis mRNA
Log10 Embryo mRNA Counts (FPKM)
Log10 Adult mRNA Counts (FPKM)
rs = 0.02
G
Total mRNA
Glycolysis mRNA
Log10 Embryo mRNA Counts (FPKM)
Log10 Adult mRNA Counts (FPKM)
rs = 0.71
H
Total mRNA
TCA cycle mRNA
Log10 Embryo mRNA Counts (FPKM)
Log10 Adult mRNA Counts (FPKM)
rs = 0.78

SUPPORTING INFORMATION

Comparison of *Drosophila melanogaster* embryo and adult proteome by SWATH-MS reveals differential regulation of protein synthesis, degradation machinery and metabolism modules

Bertrand Fabre[1,2,4,5,†,*], Dagmara Korona[3,4,†], Jonathan G. Lees[6], Ikrame Lazar[5], Ido Livneh[5], Manon Brunet[1,2,4], Christine A. Orengo[6], Steven Russell[3,4], Kathryn S. Lilley[1,2,4,*]

1. Cambridge Centre for Proteomics, Department of Biochemistry, University of Cambridge, Cambridge, U.K

2. Department of Biochemistry, University of Cambridge, University of Cambridge, Cambridge, U.K

3. Department of Genetics, University of Cambridge, University of Cambridge, Cambridge, U.K

4. Cambridge Systems Biology Centre, University of Cambridge, Cambridge, U.K

5. Technion Integrated Cancer Center (TICC), The Rappaport Faculty of Medicine and Research Institute, Haifa, Israel

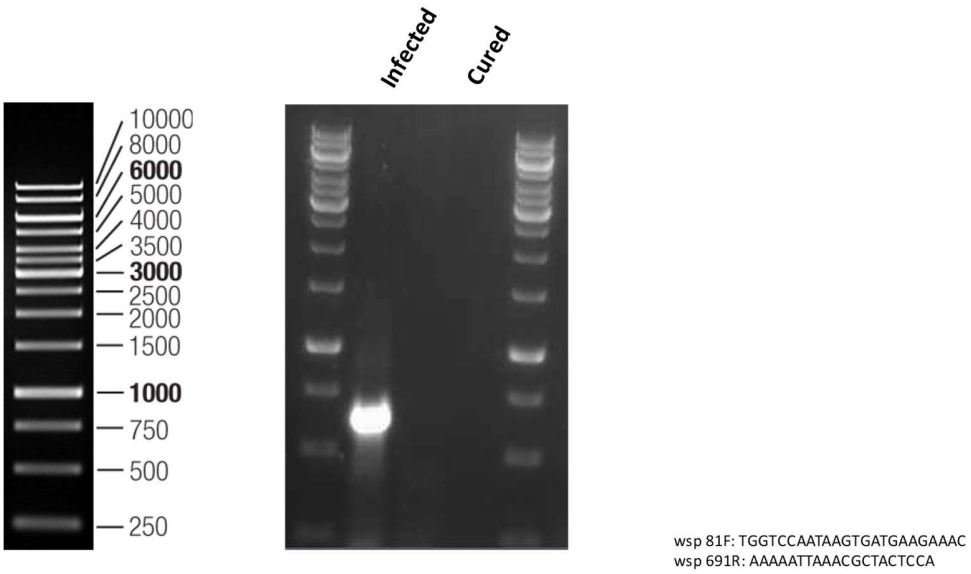
6. Institute of Structural and Molecular Biology, University College London, London, United Kingdom

† These authors contributed equally to this work.

* Correspondence:

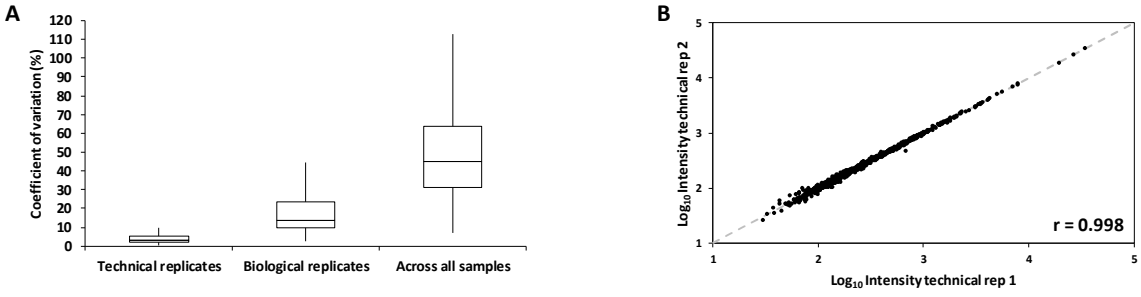
Pr. Kathryn S. Lilley, Ph.D., Cambridge Centre for Proteomics, Cambridge Systems Biology Centre, Department of Biochemistry, University of Cambridge, Cambridge CB2 1QR, U.K. Telephone: 01223 760255; e-mail: k.s.lilley@bioc.cam.ac.uk

Dr. Bertrand Fabre, Ph.D., Technion Integrated Cancer Center (TICC), The Rappaport Faculty of Medicine and Research Institute, Haifa, Israel, e-mail: bertrand.fabre@cantab.net



632 bp fragment of the Wolbachia Surface Protein gene (genbank accession no: EU395833.1).

Figure S1: Verification of Wolbachia infection status of fly lines by PCR



C

	Technical rep 1	Technical rep 2	Technical rep 3	Technical rep 4
Technical rep 1		0.998	0.998	0.997
Technical rep 2			0.998	0.997
Technical rep 3				0.997
Technical rep 4				

Figure S2: Reproducibility of the SWATH-MS workflow

A. Coefficients of variation were calculated for each protein quantified between technical (injection), biological replicates or across all the samples. B. Pairwise comparison between two technical (injection) replicates. Pearson coefficient of correlation is indicated on the graph. C. Pearson coefficient of correlation calculated from pairwise comparison of 4 technical (injection) replicates.

	Adult Wp - Rep 1	Adult Wp - Rep 2	Adult Wp - Rep 3	Adult Wp - Rep 4	Adult Wp - Rep 5	Adult Wp + Rep 1	Adult Wp + Rep 2	Adult Wp + Rep 3	Adult Wp + Rep 4	Adult Wp + Rep 5	Embryo Wp - Rep 1	Embryo Wp - Rep 2	Embryo Wp - Rep 3	Embryo Wp - Rep 4	Embryo Wp - Rep 5	Embryo Wp + Rep 1	Embryo Wp + Rep 2	Embryo Wp + Rep 3	Embryo Wp + Rep 4	Embryo Wp + Rep 5
Adult Wp - Rep 1	1.00	0.92	0.92	0.92	0.93	0.92	0.91	0.93	0.92	0.92	0.61	0.65	0.56	0.61	0.59	0.58	0.65	0.60	0.58	0.57
Adult Wp - Rep 2		1.00	0.93	0.92	0.94	0.92	0.92	0.92	0.93	0.91	0.62	0.65	0.56	0.61	0.59	0.58	0.65	0.60	0.58	0.57
Adult Wp - Rep 3			1.00	0.92	0.93	0.93	0.93	0.92	0.93	0.92	0.63	0.67	0.57	0.63	0.60	0.60	0.67	0.61	0.59	0.59
Adult Wp - Rep 4				1.00	0.92	0.92	0.92	0.91	0.92	0.92	0.62	0.66	0.56	0.62	0.59	0.58	0.66	0.60	0.58	0.57
Adult Wp - Rep 5					1.00	0.93	0.93	0.93	0.94	0.92	0.62	0.66	0.56	0.62	0.60	0.59	0.66	0.60	0.58	0.58
Adult Wp + Rep 1						1.00	0.92	0.91	0.92	0.92	0.64	0.67	0.58	0.63	0.61	0.60	0.67	0.62	0.60	0.59
Adult Wp + Rep 2							1.00	0.92	0.92	0.92	0.64	0.67	0.57	0.63	0.61	0.60	0.67	0.62	0.59	0.59
Adult Wp + Rep 3								1.00	0.93	0.91	0.63	0.67	0.57	0.63	0.61	0.60	0.67	0.62	0.60	0.59
Adult Wp + Rep 4									1.00	0.93	0.64	0.67	0.58	0.64	0.61	0.61	0.67	0.62	0.60	0.60
Adult Wp + Rep 5										1.00	0.65	0.68	0.59	0.65	0.62	0.61	0.68	0.62	0.61	0.61
Embryo Wp - Rep 1											1.00	0.96	0.93	0.97	0.96	0.95	0.96	0.97	0.97	0.97
Embryo Wp - Rep 2												1.00	0.89	0.96	0.93	0.91	0.98	0.96	0.94	0.94
Embryo Wp - Rep 3													1.00	0.92	0.96	0.97	0.90	0.94	0.96	0.94
Embryo Wp - Rep 4														1.00	0.96	0.95	0.97	0.97	0.97	0.97
Embryo Wp - Rep 5															1.00	0.97	0.94	0.96	0.97	0.97
Embryo Wp + Rep 1																1.00	0.92	0.95	0.97	0.96
Embryo Wp + Rep 2																	1.00	0.97	0.95	0.95
Embryo Wp + Rep 3																		1.00	0.97	0.97
Embryo Wp + Rep 4																			1.00	0.98
Embryo Wp + Rep 5																				1.00

Figure S3: Pearson coefficient of correlation calculated between all the samples from pairwise comparison.

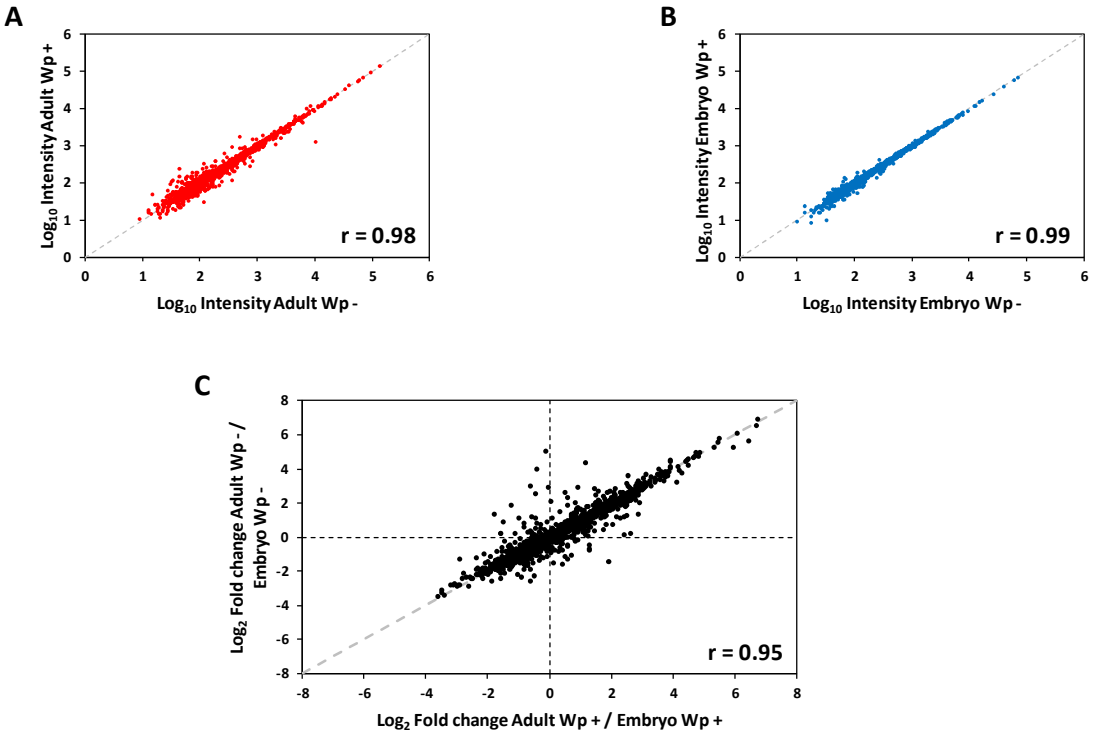


Figure S4: Comparison of Wolbachia infected and cured samples

A-B. Comparison of average protein intensities between Wolbachia infected (Wp+) and cured (Wp-) adult fly (A) and embryo (B). Pearson coefficient of correlation is indicated on the graphs.

C. Comparison of the average ratio adult/embryo protein intensities between Wolbachia infected (Wp+) and cured (Wp-) adult fly (A) and embryo (B) infected with Wolbachia (Wp+) and cured (Wp-). Pearson coefficient of correlation is indicated on the graph.

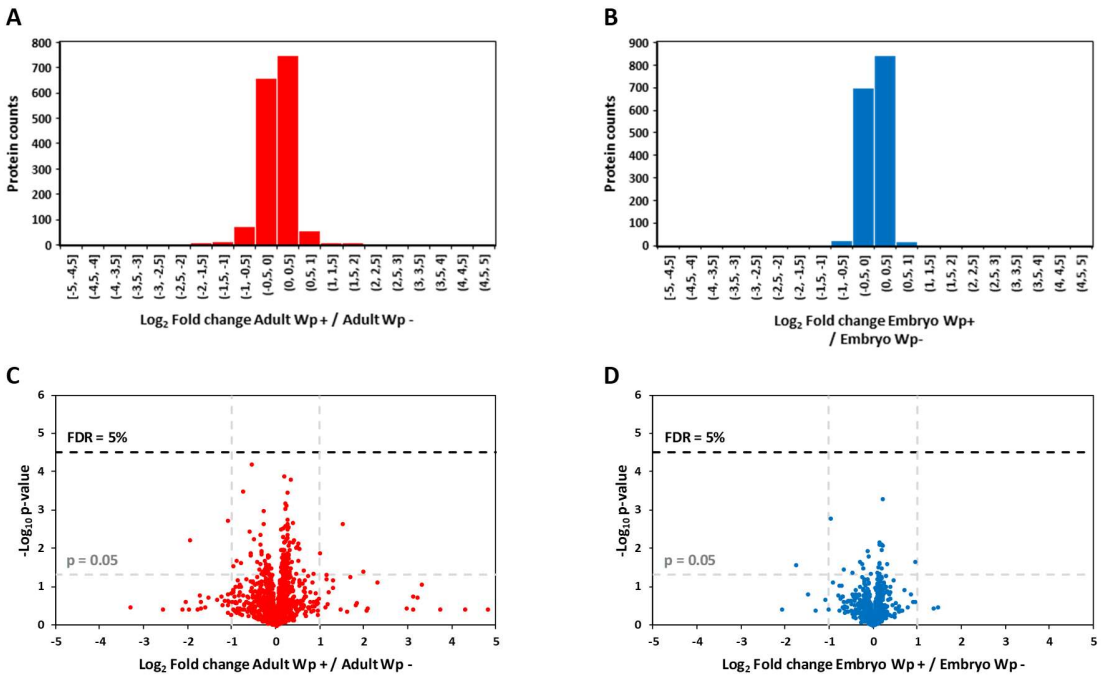


Figure S5: Effect of *Wolbachia pipientis* on the embryo and adult fly proteomes

A-B. Graphs representing the distribution of $\log_2$ ratio Wolbachia infected (Wp+) / cured (Wp-) for adult fly (A) and embryo (B). C-D. Volcano plot representing the $\log_2$ ratio Wolbachia infected (Wp+) / cured (Wp-) for adult fly (C) and embryo (D) and the corresponding Benjamini-Hochberg corrected p-value. Dashed lines represent the thresholds for $\log_2$ fold change of 1 and -1 as well as p-value of 0.05 and FDR of 5%.

Figure S6A

Aconitate (Q9VIE8_DROME) Spectronaut average measured Adult/Embryo ratio : 11.4

WVAVGDENYGEGSSR

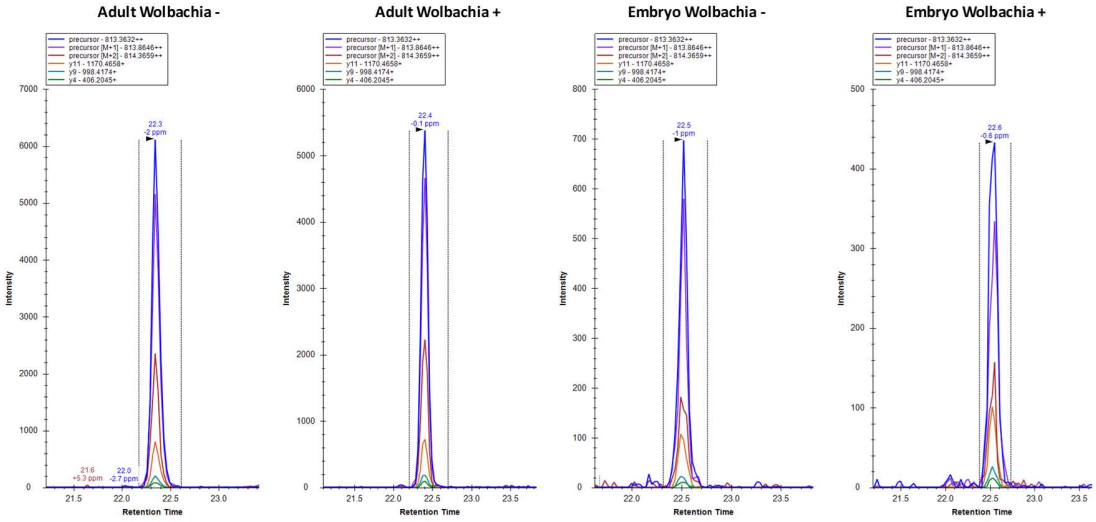


Figure S6B

ATP synthase subunit alpha (Protein bellwether)(ATPA_DROME) Spectronaut average measured Adult/Embryo ratio : 5.8

QGQYVPMAIEDQVAVIYCGVR

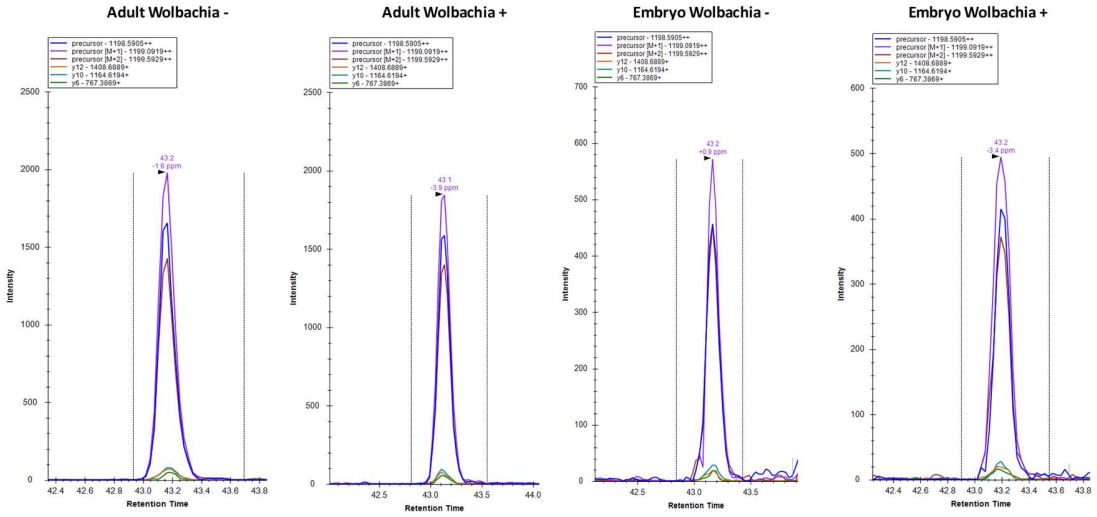


Figure S6C

Proteasome subunit alpha type-1 (PSA1_DROME) Spectronaut average measured Adult/Embryo ratio : 0.33

HSYDTTYPVSR

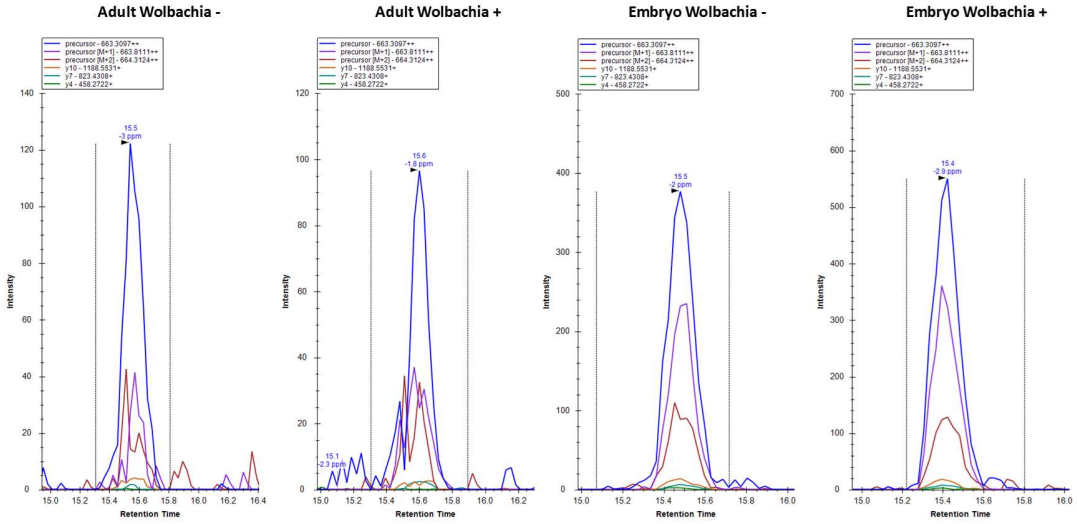


Figure S6D

60S ribosomal protein L4 (RL4_DROME) Spectronaut average measured Adult/Embryo ratio : 0.38

FVIWTESAFAR

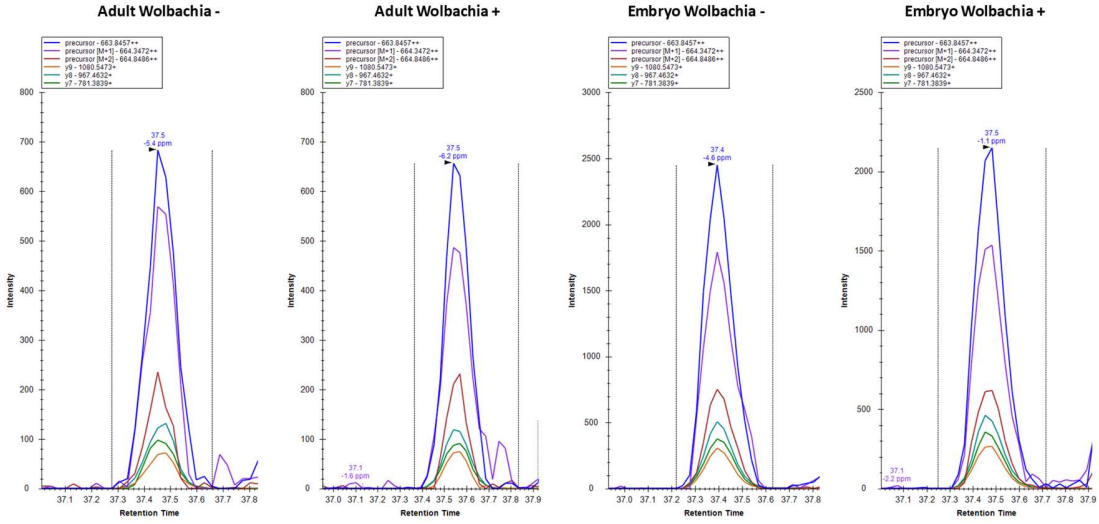


Figure S6: Skyline validation of the SWATH-MS data

A-D. Extracted ion chromatograms from the re-analysis of the data using Skyline are presented for peptides of Aconitate (A), ATP synthase subunit alpha (Protein bellwether) (B), Proteasome subunit alpha type-1 (C) and 60S ribosomal protein L4 (D). The adult/embryo ratio measured using Spectronaut is given for each protein.

Figure S7A

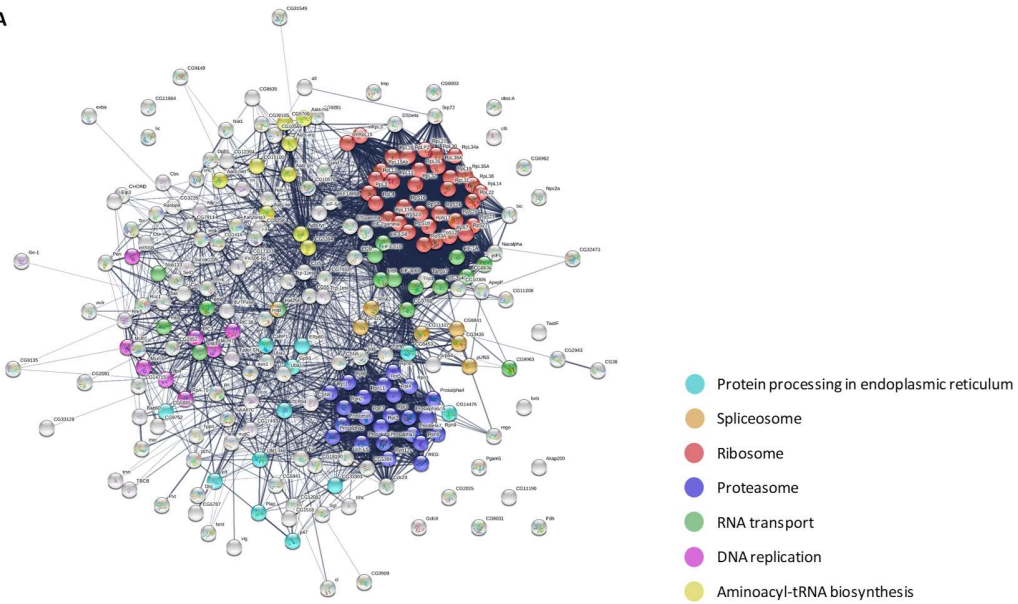


Figure S7B

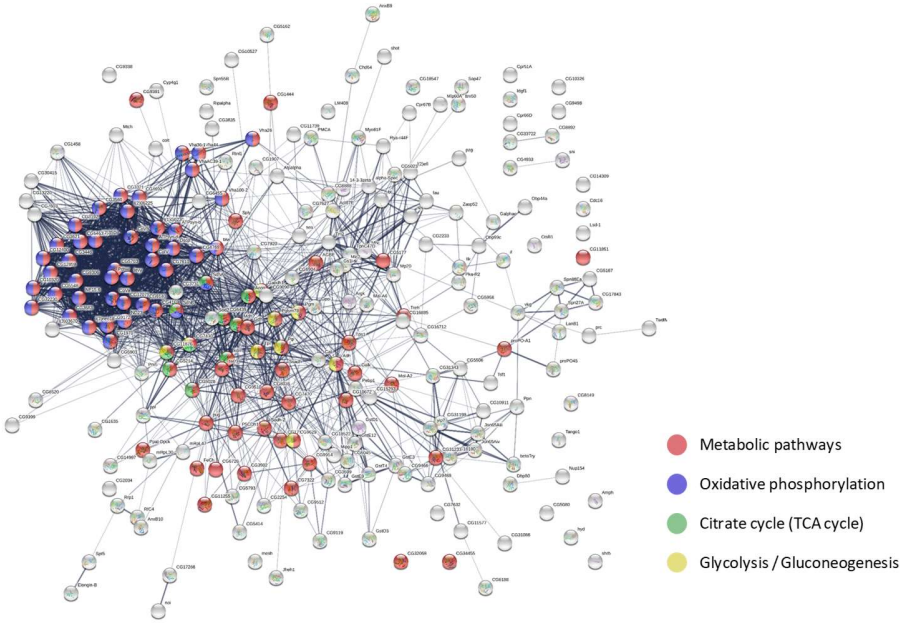


Figure S7: Network representation of differentially regulated proteins between adult fly and embryo

A-B. Graph representing the STRING analysis of the proteins more abundant in embryo (A) and adult flies (B). Proteins belonging to significantly enriched pathways are highlighted.

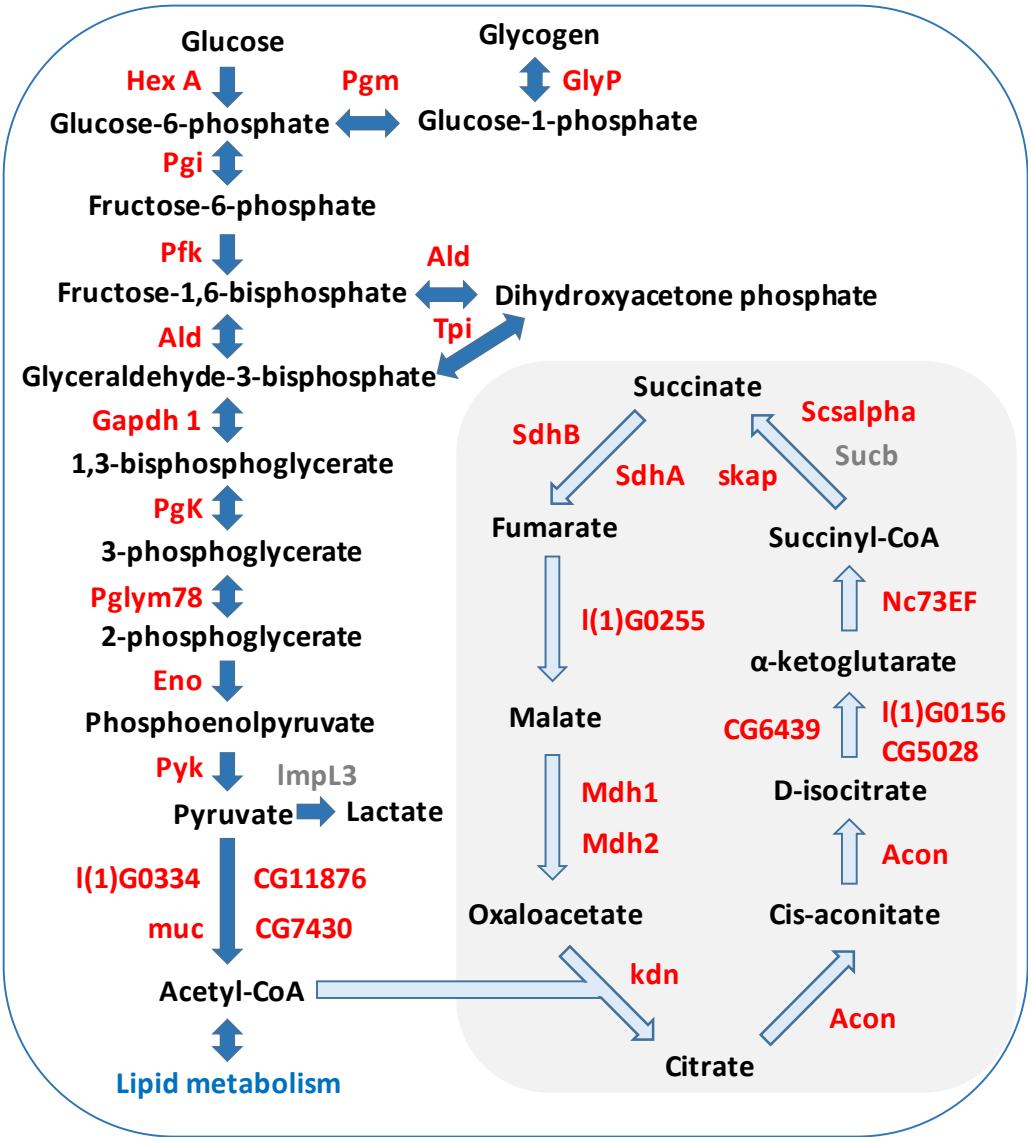


Figure S8: Proteins from glycolysis and TCA cycle are more abundant in adult fly compared to embryo

Graph displaying the enzymes involved in glycolysis and TCA cycle. Enzymes with name in blue, red and grey represent proteins more abundant in embryo, in adult flies and not significantly changing, respectively.

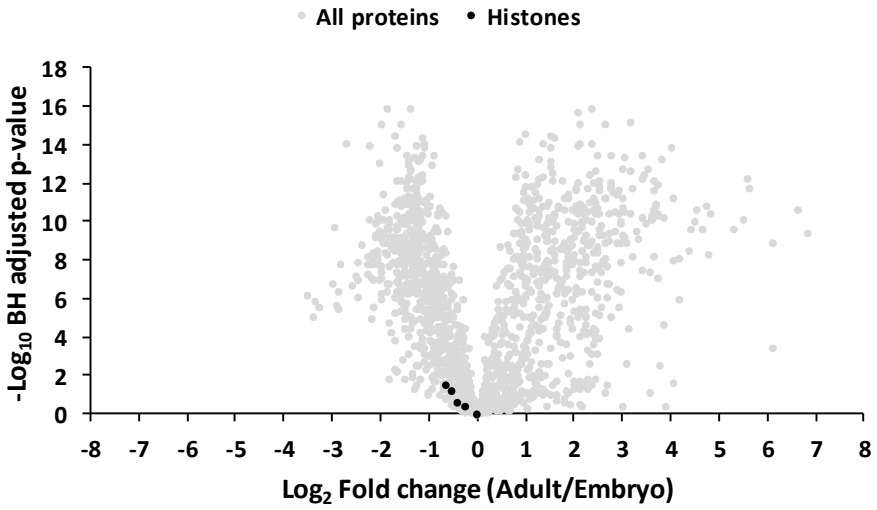


Figure S9: Histones remain constant between embryo and adult fly

Volcano plot representing the $\log_2$ ratio (adult/embryo) for each protein quantified and the corresponding Benjamini-Hochberg corrected p-value. Histones are represented of the graphs.

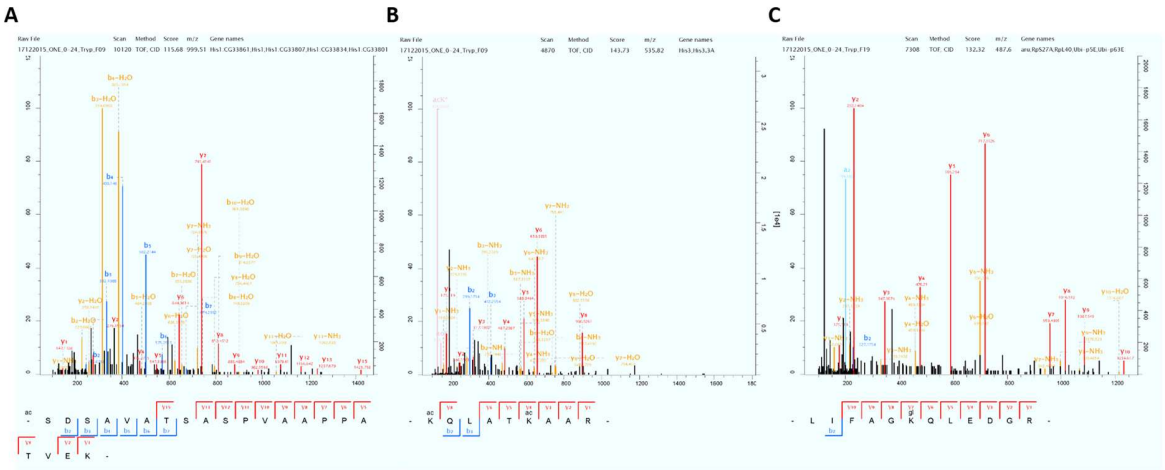


Figure S10: MSMS spectra for acetylated histones and K48 ubiquitin linkage

A-C. MSMS spectra of the acetylated peptide acSDSAVATSASPVAAPPATVEK of the histone H1 (A), the acetylated peptide acKQLATacKAAR of the histone H3 (B) and the peptide reporting K48 ubiquitin linkage LIFAGK[GG]QLEDGR (C) identified in the spectral library using MaxQuant.

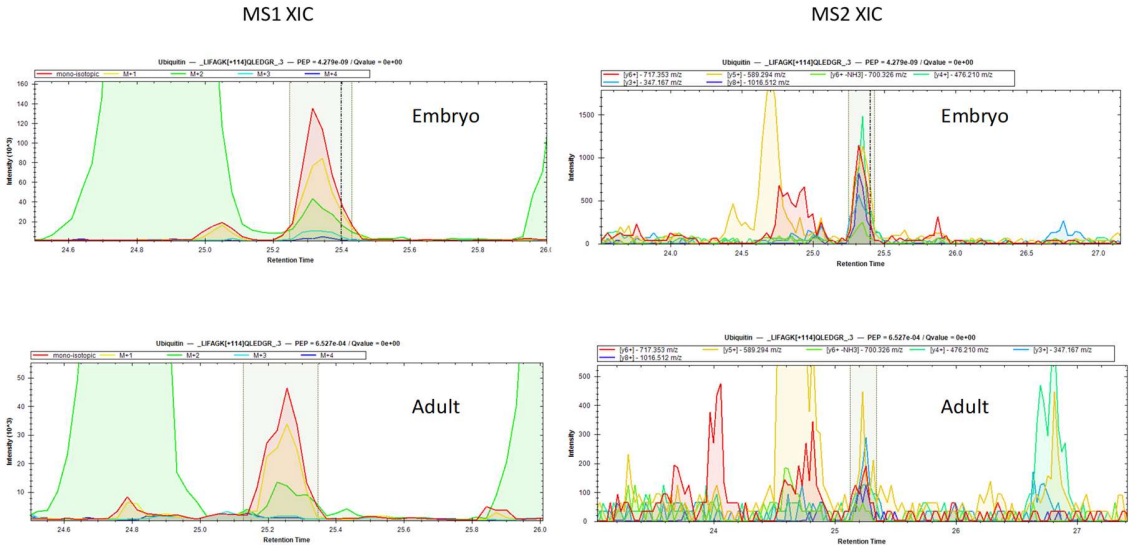


Figure S11: Extracted ion chromatograms of the peptide reporting K48 ubiquitin linkage LIFAGK[GG]QLEDGR

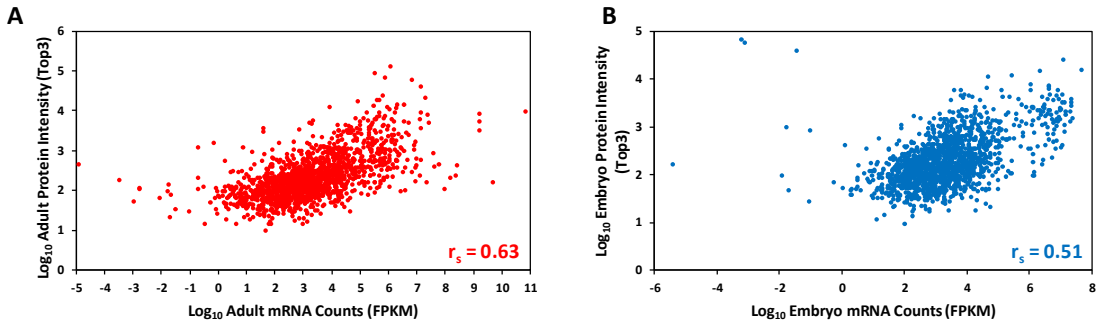


Figure S12: Protein/mRNA correlation in adult fly and embryo

A-B. Graphs representing the $\log_{10}$ protein intensities (calculated using the Top3 method) against the adult $\log_{10}$ mRNA counts (calculated using the modENCODE data) for adult fly (A) and embryo (B). The resulting Spearman's rank correlation coefficients are indicated on the graphs.

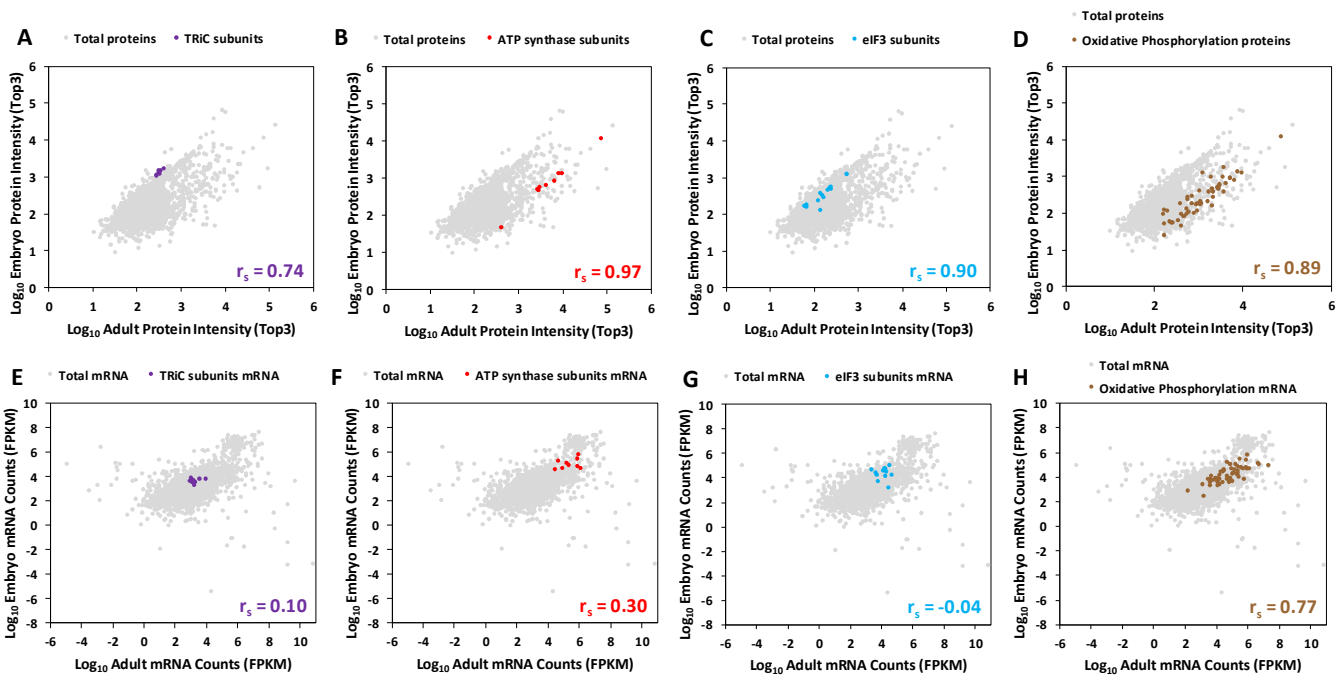


Figure S13: Correlation of Protein and mRNA components of modules

A-D. Graph representing the embryo $\log_{10}$ protein intensities (calculated using the Top3 method) against the adult $\log_{10}$ protein intensities. TRiC (A), ATP synthase (B), eIF3 (C) and Oxidative phosphorylation (D) components proteins intensities are represented on the graphs. E-H. Graph representing the embryo $\log_{10}$ mRNA counts (calculated using the modENCODE data) against the adult $\log_{10}$ mRNA counts. TRiC (E), ATP synthase (F), eIF3 (G) and Oxidative phosphorylation (H) components mRNA counts are represented on the graphs. The resulting Spearman's rank correlation coefficient is indicated on the graphs.

Figure S14A

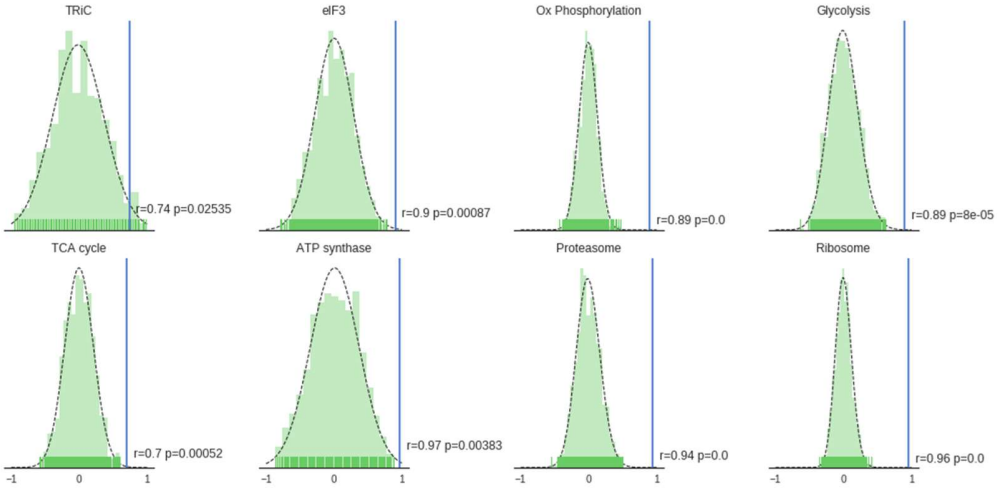


Figure S14B

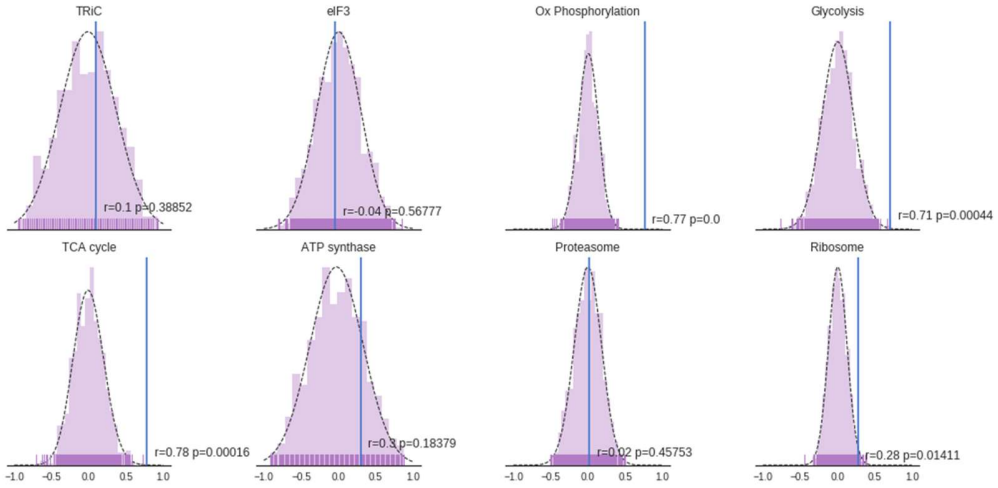


Figure S14: Random sampling experiments with the sample size matching the number of genes quantified in each protein module

A-B. Graph representing the skew-normal distribution fitted to the Spearman's rank correlation coefficient distribution of 1,000 random samples at the protein (A) and mRNA (B) level. The corresponding p-value was estimated for each protein module.

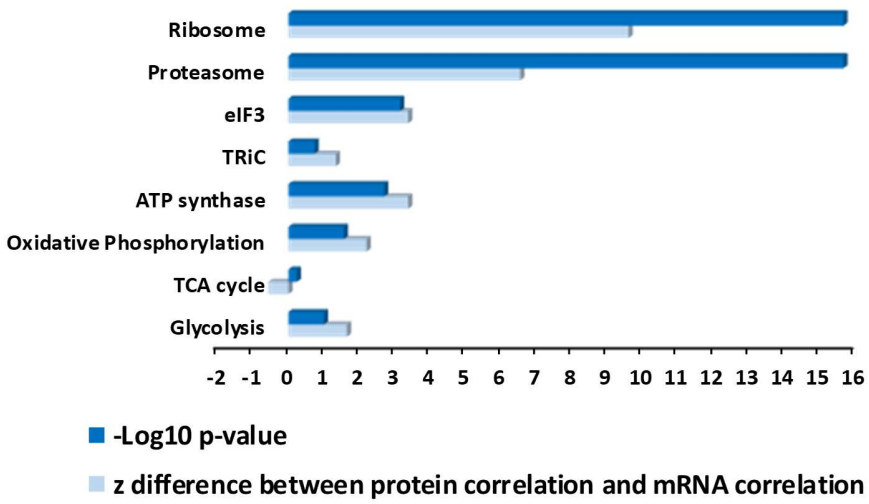


Figure S15: Differences between Protein and mRNA components of modules correlation

Coefficient of correlations observed at the protein and mRNA level for these protein modules were transformed using Fisher r-to-z transformation to assess the difference between the correlations. A graph representing the z differences and the corresponding $-\log_{10}$ p-values for the different protein modules is displayed.

Table S1. Table containing the SWATH-MS data

Table S2. Table containing the protein and mRNA data

Table S3. Table containing the post-translational modifications data